#### WORLD INTELLECTUAL PROPERTY ORGANIZATION International Bureau



## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7: C12N 15/86, 5/10, 7/04, 15/49, C07K **A1** 14/16

(11) International Publication Number:

WO 00/15819

(43) International Publication Date:

23 March 2000 (23:03:00)

(21) International Application Number:

PCT/US99/20675

(22) International Filing Date:

10 September 1999 (10.09.99)

(30) Priority Data:

60/100,022

11 September 1998 (11.09.98)

60/100,063

US 12 September 1998 (12.09.98)

(71) Applicant: THE CHILDREN'S MEDICAL CENTER COR-PORATION [US/US]; 300 Longwood Avenue, Boston, MA 02115 (US).

(72) Inventors: GRAY, John, T.; 48 Spring Road, West Roxbury, MA 02132 (US). MULLIGAN, Richard, C.; 2 Sandy Pond Road, Lincoln, MA 01773 (US).

(74) Agents: BROOK, David, E. et al.; Hamilton, Brook, Smith & Reynolds, P.C., Two Militia Drive, Lexington, MA 02421 (US).

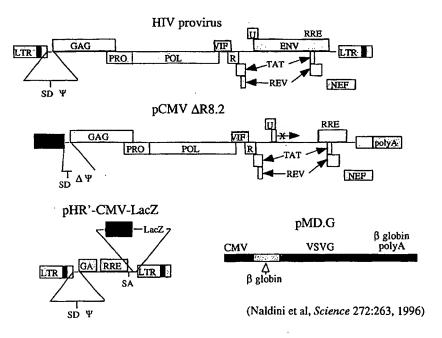
(81) Designated States: AU, CA, JP, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

#### Published

With international search report.

Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: PACKAGING CELL LINES FOR HIV-DERIVED RETROVIRAL VECTOR PARTICLES



#### (57) Abstract

Novel packaging cell lines useful for generating viral accessory protein independent HIV-derived retroviral vector particles, methods of constructing such packaging cell lines and methods of using the viral accessory protein independent HIV-derived retroviral vector particles are disclosed.

## FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL AM AT AU AZ BA BB BE BF BG BJ BR BY CA CF CG CH CI CM CN CU CZ DE DK EE	Albania Armenia Australia Australia Azerbaijan Bosnia and Herzegovina Barbados Belgium Burkina Faso Bulgaria Benin Brazil Belarus Canada Central African Republic Congo Switzerland Côte d'Ivoire Cameroon China Cuba Czech Republic Germany Denmark Estonia	ES FI FR GA GB GE GH GN GR HU IE IL IS IT JP KE KG KP  KR LC LI LK LR	Spain Finland France Gabon United Kingdom Georgia Ghana Guinea Greece Hungary Ireland Israel Iceland Italy Japan Kenya Kyrgyzstan Democratic People's Republic of Korea Republic of Korea Kazakstan Saint Lucia Liechtenstein Sri Lanka Liberia	LS LT LU LV MC MD MG MK ML MN MR MW MX NE NL NO NZ PL PT RO RU SD SE SG	Lesotho Lithuania Luxembourg Latvia Monaco Republic of Moldova Madagascar The former Yugoslav Republic of Macedonia Mali Mongolia Mauritania Malawi Mexico Niger Netherlands Norway New Zealand Poland Portugal Romania Russian Federation Sudan Sweden Singapore	SI SK SN SZ TD TG TJ TM TR TT UA UG US UZ VN YU ZW	Slovenia Slovakia Senegal Swaziland Chad Togo Tajikistan Turkmenistan Turkey Trinidad and Tobago Ukraine Uganda United States of America Uzbekistan Viet Nam Yugoslavia Zimbabwe
--	--	---	---	---	---	--	--

-1-

## PACKAGING CELL LINES FOR HIV-DERIVED RETROVIRAL VECTOR PARTICLES

#### BACKGROUND OF THE INVENTION

Retroviral vectors based on lentiviruses, such as human immunodeficiency viruses (HIV), can infect nondividing cells, and integration of proviral DNA occurs without the need for cell division. These properties make lentiviruses attractive for gene transfer into nondividing cells, such as hepatocytes, myofibers, hematopoietic stem cells, and neurons.

However, the use of lentivirus vectors, particularly HIV vectors, particularly for gene therapy, is hampered by concern over their safety. Thus, a need for the development of lentivirus vectors, particularly HIV vectors, with improved safety, particularly for gene therapy, exists.

#### SUMMARY OF THE INVENTION

15

20

The present invention relates to novel packaging cell lines useful for generating viral accessory protein independent lentivirus-derived, particularly HIV-derived, retroviral vector particles, to construction of such cell lines and to methods of using the accessory protein independent lentivirus-derived retroviral vector particles to introduce DNA of interest into cells (e.g, eukaryotic cells such as animal (particularly mammalian), plant or yeast cells or prokaryotic cells such as bacterial cells). In a preferred embodiment, the packaging cell lines of the present invention are stable packaging cell lines.

In one embodiment of the invention, packaging cell lines for producing a viral accessory protein independent lentivirus-derived retroviral vector particles comprise (a) a cell (e.g., mammalian cell); and (b) a retroviral nucleotide sequence in the cell which comprises a coding sequence for lentivirus gappol, wherein said coding sequence has

been codon optimized by mutagenisis to improve expression of the lentivirus gagpol proteins.

In second embodiment of the invention, packaging cell lines for producing a viral accessory protein independent lentivirus-derived retroviral vector particles comprise (a) a cell (e.g., mammalian cell); (b) a first retroviral nucleotide sequence in the cell which comprises a coding sequence for lentivirus gagpol, wherein said coding sequence has been codon optimized by mutagenisis to improve expression of the lentivirus gagpol proteins; and (c) a second retroviral nucleotide sequence in the cell which comprises the coding sequence for a heterologous envelope protein.

10

15

20

25

In a third embodiment of the invention, packaging cell lines for producing a viral accessory protein independent lentivirus-derived retroviral vector particles comprise (a) a cell (e.g., mammalian cell); (b) a first retroviral nucleotide sequence in the cell which comprises a coding sequence for lentivirus gagpol, wherein said coding sequence has been codon optimized by mutagenisis to improve expression of the lentivirus gagpol proteins; (c) a second retroviral nucleotide sequence in the cell which comprises the coding sequence for a heterologous envelope protein; and (d) a third retroviral nucleotide sequence which comprises a DNA sequence of interest and lentivirus cisacting sequences required for packaging, reverse transcription and integration.

In a fourth embodiment of the invention, packaging cell lines for producing a viral accessory protein independent lentivirus-derived retroviral vector particles comprise (a) a cell (e.g., mammalian cell); (b) a retroviral nucleotide sequence in the cell which comprises a coding sequence for lentivirus gagpol, wherein said coding sequence has been codon optimized by mutagenisis to improve expression of the lentivirus gagpol proteins; and (c) a retroviral nucleotide sequence which comprises a DNA sequence of interest and lentivirus cis-acting sequences required for packaging, reverse transcription and integration.

In a fifth embodiment of the invention, packaging cell lines for producing a viral accessory protein independent HIV-derived retroviral vector particles comprise (a) a cell

(e.g., mammalian cell); and (b) a retroviral nucleotide sequence in the cell which comprises a coding sequence for HIV *gagpol*, wherein said coding sequence has been codon optimized by mutagenisis to improve expression of the HIV gagpol proteins.

In sixth embodiment of the invention, packaging cell lines for producing a viral accessory protein independent HIV-derived retroviral vector particles comprise (a) a cell (e.g., mammalian cell); (b) a first retroviral nucleotide sequence in the cell which comprises a coding sequence for HIV gagpol, wherein said coding sequence has been codon optimized by mutagenisis to improve expression of the HIV gagpol proteins; and (c) a second retroviral nucleotide sequence in the cell which comprises the coding sequence for a heterologous envelope protein.

10

20

In a seventh embodiment of the invention, packaging cell lines for producing a viral accessory protein independent HIV-derived retroviral vector particles comprise (a) a cell (e.g., mammalian cell); (b) a first retroviral nucleotide sequence in the cell which comprises a coding sequence for HIV *gagpol*, wherein said coding sequence has been codon optimized by mutagenisis to improve expression of the HIV gagpol proteins; (c) a second retroviral nucleotide sequence in the cell which comprises the coding sequence for a heterologous envelope protein; and (d) a third retroviral nucleotide sequence which comprises a DNA sequence of interest and HIV cis-acting sequences required for packaging, reverse transcription and integration.

In a eighth embodiment of the invention, packaging cell lines for producing a viral accessory protein independent HIV-derived retroviral vector particles comprise (a) a cell (e.g., mammalian cell); (b) a retroviral nucleotide sequence in the cell which comprises a coding sequence for HIV gagpol, wherein said coding sequence has been codon optimized by mutagenisis to improve expression of the HIV gagpol proteins; and (c) a retroviral nucleotide sequence which comprises a DNA sequence of interest and HIV cis-acting sequences required for packaging, reverse transcription and integration.

Alternatively, each of the packaging cell lines described herein can be produced using (1) a retroviral nucleotide sequence which comprises a codon optimized gag

coding sequence and (2) a retroviral nucleotide sequence which comprises a codon optimized pol coding sequence, in place of the retroviral nucleotide sequence which comprises a codon optimized gagpol coding sequence.

In a particular embodiment, the heterologous envelope protein is the G glycoprotein of vesicular stomatitis virus (VSV G). In another embodiment, the heterologous envelope protein is the amphotropic envelope of the Moloney leukemia virus (MLV).

5

10

20

Cell lines for producing a viral accessory protein independent lentivirus-derived retroviral vector particles are produced by transfecting host cells (e.g., mammalian host cells) with a plasmid comprising a DNA sequence which encodes lentivirus gagpol proteins, wherein said DNA sequence has been codon optimized by mutagenisis to improve expression of the lentivirus gagpol proteins. Depending upon the particular cell line being produced, the host cells are also co-transfected with a plasmid comprising a DNA sequence which encodes a heterologous envelope protein, or a plasmid comprising a DNA sequence of interest and lentivirus cis-acting sequences required for packaging, reverse transcription and integration, or both of these plasmids. Alternatively, host cells are transfected with a plasmid comprising a codon optimized DNA sequence encoding a lentivirus gag protein and a plasmid comprising a codon optimized DNA sequence encoding a lentivirus pol protein, in place of the plasmid comprising a codon optimized DNA sequence encoding both lentivirus gagpol proteins.

Cell lines for producing a viral accessory protein independent HIV-derived retroviral vector particles are produced by co-transfecting host cells (e.g., mammalian host cells) with a plasmid comprising a DNA sequence which encodes HIV gagpol proteins, wherein said DNA sequence has been codon optimized by mutagenisis to improve expression of the HIV gagpol proteins. Depending upon the particular cell line being produced, the host cells are also co-transfected with a plasmid comprising a DNA sequence which encodes a heterologous envelope protein, or a plasmid comprising a DNA sequence of interest and HIV cis-acting sequences required for packaging, reverse

-5-

transcription and integration, or both of these plasmids. Alternatively, host cells are transfected with a plasmid comprising a codon optimized DNA sequence encoding a HIV gag protein and a plasmid comprising a codon optimized DNA sequence encoding a HIV pol protein, in place of the plasmid comprising a codon optimized DNA sequence encoding both HIV gagpol proteins.

The present invention also relates to methods of producing viral accessory protein independent lentivirus-derived retroviral vector particles, comprising cotransfecting host cells (e.g., mammalian host cells) with (a) a first plasmid comprising a DNA sequence which encodes lentivirus gagpol proteins, wherein said DNA sequence has been codon optimized by mutagenisis to improve expression of the lentivirus gagpol proteins; (b) a second plasmid comprising a DNA sequence which encodes a heterologous envelope protein; and (c) a third plasmid comprising a DNA sequence of interest and lentivirus cis-acting sequences required for packaging, reverse transcription and integration. Alternatively, host cells are transfected with a plasmid comprising a codon optimized DNA sequence encoding a lentivirus gag protein and a plasmid comprising a codon optimized DNA sequence encoding a lentivirus pol protein, in place of the first plasmid comprising a codon optimized DNA sequence encoding both lentivirus gagpol proteins.

10

15

In a particular embodiment, the invention relates to methods of producing viral
accessory protein independent HIV-derived retroviral vector particles, comprising cotransfecting host cells (e.g., mammalian host cells) with (a) a first plasmid comprising a
DNA sequence which encodes HIV gagpol proteins, wherein said DNA sequence has
been codon optimized by mutagenisis to improve expression of the HIV gagpol
proteins; (b) a second plasmid comprising a DNA sequence which encodes a
heterologous envelope protein; and (c) a third plasmid comprising a DNA sequence of
interest and HIV cis-acting sequences required for packaging, reverse transcription and
integration. Alternatively, host cells are transfected with a plasmid comprising a codon
optimized DNA sequence encoding a HIV gag protein and a plasmid comprising a

PCT/US99/20675

codon optimized DNA sequence encoding a HIV pol protein, in place of the first plasmid comprising a codon optimized DNA sequence encoding both HIV gagpol proteins.

The present invention also relates to viral accessory protein-independent retroviral particles produced by or obtainable by (obtained by) the methods described herein.

The present invention further relates to isolated DNA encoding a codon optimized lentivirus gagpol, isolated DNA encoding the gag coding region of a codon optimized lentivirus gagpol, and isolated DNA encoding the pol coding region of a codon optimized lentivirus gagpol. In a particular embodiment, the present invention relates to isolated DNA encoding a codon optimized HIV gagpol, isolated DNA encoding the gag coding region of a codon optimized HIV gagpol, and isolated DNA encoding the pol coding region of a codon optimized HIV gagpol.

The packaging cell lines and viral particles of the present invention can be used for gene therapy or gene replacement with improved safety. The packaging cell lines and viral particles of the present invention can also be used in development and production of vaccines, and in production of biochemical reagents. Gene therapy vectors produced with the cell lines of the present invention are expected to be valuable medical therapeutics.

# 20 BRIEF DESCRIPTION OF THE DRAWINGS

15

Figure 1 is a schematic diagram of an expression cassette containing the codon optimized gagpol genes. The DNA was constructed in multiple segments, which are indicated at the top as 1/3, 2/3, 3/3 (A, B, C and D) and HIN. Restriction sites used to assemble the cloned segments are indicated above the kilobasepair (Kb) ruler. Below the ruler are multiple features showing the location of the human cytomegalovirus (CMV) promoter, human betaglobin sequences (Bglobin), mRNA sequences (thinner line represents intronic sequence), the gag and pol open reading frames, the individual

proteolytic fragment coding sequences (p17\_MA, p24\_CA, p7, p6, PR, p51\_RT, RNaseH and integrase (IN)) and each synthetic oligonucleotide used in the assembly process (multiple adjacent open arrows).

Figure 2 is a table which depicts codon usage frequencies in genes which are highly expressed and in the codon optimized gagpol open reading frame of the HIV packaging construct described herein.

Figure 3 is a schematic representation of the HIV provirus and a three-plasmid expression system used for generating a pseudotyped HIV-based vector by transient transfection as described in Naldini *et al.*, *Science*, *272*:263-267 (1996).

Figure 4 is a list of some characteristics relating to the HIV Rev protein.

Figure 5 is a list of some points relating to codon optimization of HIV gagpol.

Figure 6 is a partial DNA sequence of HIV gag (SEQ ID NO: 1), showing inactivation of inhibitory sequences as described in Schwartz, S. et al., J. Virol., 66(12):7176-7182 (1992).

10

- 15

20

25

Figure 7 a plot of the %(G+C) content of wildtype HIV *gagpol* sequences and theoretically codon optimized HIV *gagpol* sequences. The percent of bases, either G or C, was calculated for a 30 nucleotide moving window for the entire length of the *gagpol* gene, and the value plotted versus nucleotide position. Diamonds = HIV *gagpol* sequences; squares = full optimal back-translation for *gag* open reading frame; triangles = full optimal back-translation for *pol* open reading frame; CO = codon optimized.

Figures 8A-8E depict the alignment of the nucleotide sequences and predicted amino acid sequences for the *gag* coding region of a wildtype HIV *gagpol* and a codon optimized HIV *gagpol*. "NL4-3 genbank.SEQ" indicates the nucleotide sequence (SEQ ID NO:2) and predicted amino acid sequence (SEQ ID NO:3) for the *gag* coding region of a wildtype HIV *gagpol*. "pHDMHgpm2.seq" indicates the nucleotide sequence (SEQ ID NO:4) and predicted amino acid sequence (SEQ ID NO:5) for the *gag* coding region

of a codon optimized HIV gagpol. The "NL4-3 genbank.SEQ" sequences are publicly available at the NIH GenBank sequence repository (Accesssion No. M19921).

Figures 9A-9L depict the alignment of the nucleotide sequences and predicted amino acid sequences for the *pol* coding region of a wildtype HIV *gagpol* and a codom optimized HIV *gagpol*. "NL4-3 genbank.SEQ" indicates a nucleotide sequence (SEQ ID NO:6) and a predicted amino acid sequence (SEQ ID NO:7) for the *pol* coding region of a wildtype HIV *gagpol* available in the NIH GenBank sequence repository (Accesssion No. M19921). The nucleotide and amino acid sequences for the *pol* coding region available in the GenBank sequence repository contain two sequence errors, which are indicated in Figures 9A-9L with shading. "pNL4-3.seq" indicates the correct nucleotide sequence (SEQ ID NO:8) and predicted amino acid sequence (SEQ ID NO:9) for the *pol* coding region of a wildtype HIV *gagpol*. "pHDMHgpm2.seq" indicates the nucleotide sequence (SEQ ID NO:10) and predicted amino acid sequence (SEQ ID NO:11) for the *pol* coding region of a codon optimized HIV *gagpol*.

Figures 10A-10D depict the DNA sequence (SEQ ID NO:12) for pHDMHgpm2. The CMV enhancer/promoter is at nucleotides 97 to 679, human betaglobin sequences (Bglobin) are at nucleotides 761 to 864, 865 to 1303 and 5710 to 6469 (end of Bglobin is at nucleotides 6445 to 6469), mRNA sequences are at nucleotides 680 to 778 and 1255 to 5921, SV40 origin of replication is at nucleotides 8796 to 8908, beta-lactamase (bla) coding region is at nucleotides 6709 to 7569, intron sequences are at nucleotides 779 to 1254, the codon optimized *gag* coding region is at nucleotides 1318 to 2820, the codon optimized *pol* coding region is at nucleotides 2619 to 5624 and the poly A site is at nucleotides 5897 to 5921.

Figure 11 is a circular map of plasmid pHDMHgpm2.

## 25 DETAILED DESCRIPTION OF THE INVENTION

The present invention relates to novel packaging cell lines useful for generating viral accessory protein independent lentivirus-derived, particularly HIV-derived,

- 15

retroviral vector particles, to construction of such cell lines and to methods of using the accessory protein independent lentivirus-derived retroviral vector particles to introduce DNA of interest into cells (e.g, eukaryotic cells such as animal (particularly mammalian), plant or yeast cells or prokaryotic cells such as bacterial cells). In a particular embodiment, the packaging cell lines of the present invention are stable packaging cell lines.

The cell lines are engineered to express the lentivirus proteins necessary for virus particle formation (gagpol proteins), without containing DNA sequences from lentivirus accessory proteins (tat, vif, vpr, vpu, nef and rev proteins and Rev response element (RRE)). Additionally, no viral sequences (such as cis-acting elements termed constitutive transport elements (CTEs)) will be expressed as RNA of any kind. DNA sequences for lentivirus gagpol are codon optimized by extensively mutagenizing the sequences to improve expression and to reduce the risk of recombination between transfer vector sequences and gagpol messenger RNA. This greatly improves the safety of virus preparations generated from these cell lines. In a particular embodiment, the DNA sequences for lentivirus gagpol are not codon optimized in the overlap region between the gag and pol sequences and in cis-acting signals necessary for translation of pol.

Examples of lentiviruses include human immunodeficiency viruses (e.g., HIV-1, HIV-3), bovine lentiviruses (e.g., bovine immunodeficiency viruses, bovine immunodeficiency-like viruses, Jembrana disease viruses), equine lentiviruses (e.g., equine infectious anemia viruses), feline lentiviruses (e.g., feline immunodeficiency viruses, panther lentiviruses, puma lentiviruses), ovine/caprine lentiviruses (e.g., Brazilian caprine lentiviruses, caprine arthritis-encephalitis viruses, Maedi-Visna viruses, Maedi-Visna-like viruses, Maedi-Visna-related viruses, ovine lentiviruses, Visna lentiviruses), Simian AIDS retroviruses (e.g., human T-cell lymphotropic virus type 4), simian immunodeficiency viruses, simian-human immunodeficiency viruses, human lymphotrophic viruses (e.g., type III), simian T-cell lymphotrophic viruses.

In another embodiment, cell lines are engineered to express the HIV proteins necessary for virus particle formation (gagpol proteins), without containing DNA sequences from HIV accessory proteins (tat, vif, vpr, vpu, nef and rev proteins and Rev response element (RRE)). Additionally, no viral sequences (such as cis-acting elements termed constitutive transport elements (CTEs)) will be expressed as RNA of any kind. DNA sequences for a HIV gagpol are codon optimized by mutagenesis to improve expression and to reduce the risk of recombination between transfer vector sequences and gagpol messenger RNA. In a particular embodiment, the DNA sequences for HIV gagpol are not codon optimized in the overlap region between the gag and pol sequences and in cis-acting signals necessary for translation of pol.

10

20

Alternatively, each of the packaging cell lines described herein can be produced using (1) a nucleotide sequence which comprises a codon optimized gag coding sequence and (2) a nucleotide sequence which comprises a codon optimized pol coding sequence, in place of the nucleotide sequence which comprises a codon optimized gagpol coding sequence. In this embodiment, the gag and pol coding sequences can be completely codon optimized

Benefits of the present invention include the removal of potentially harmful lentivirus accessory proteins and other viral sequences, and the reduction of the risk of recombination to produce replication competent virus.

Packaging cell lines for producing a viral accessory protein independent lentivirus-derived retroviral vector particles comprise a mammalian cell and a retroviral nucleotide sequence comprising a coding sequence for a lentivirus *gagpol* which has been codon optimized. In a particular embodiment the packaging cell lines further comprise a retroviral nucleotide sequence comprising a coding sequence for a heterologous envelope protein. In a second embodiment, the packaging cell lines further comprise a retroviral nucleotide sequence comprising a coding sequence for a heterologous envelope protein and a retroviral nucleotide sequence which comprises a DNA sequence of interest and HIV cis-acting sequences required for packaging, reverse

transcription and integration. In third embodiment, the packaging cell lines further comprise a retroviral nucleotide sequence which comprises a DNA sequence of interest and HIV cis-acting sequences required for packaging, reverse transcription and integration. Alternatively, the packaging cell lines of the present invention comprise a retroviral nucleotide sequence which comprises a codon optimized gag coding sequence and (2) a retroviral nucleotide sequence which comprises a codon optimized pol coding sequence, in place of the retroviral nucleotide sequence which comprises a codon optimized gagpol coding sequence.

10

45

20

25

The coding sequence(s) for lentivirus gagpol which has (have) been codon optimized results in improved expression of the lentivirus gagpol proteins and reduces the risk of recombination between the transfer vector and gagpol messenger RNA. Codon optimization of the coding sequence(s) for lentivirus gagpol was obtained by mutagenizing for each particular amino acid residue, specific nucleic acid bases in a codon for the particular amino acid residue to a nucleic acid base which is present in a codon which occurs at a high frequency in genes which are highly expressed for the same amino acid residue. In a particular embodiment, the resulting optimized codon also does not cause introduction of mRNA splicing signals into the codon optimized sequence. Thus, in a particular embodiment, codon optimization of the coding sequence(s) for lentivirus gagpol is obtained by mutagenizing for each particular amino acid residue, specific nucleic acid bases in a codon for the particular amino acid residue to a nucleic acid base that is present in a codon which (1) occurs at a high frequency in genes which are highly expressed for the same amino acid residue and (2) does not cause introduction of mRNA splicing signals into the codon optimized sequence. Codon optimization typically results in the removal of nucleic acid base A-rich instability elements.

In a particular embodiment, the coding sequence for a HIV gagpol (pNL4-3; available through the AIDS repository, NIH; Adachi et al., J. Virol., 59:284-291 (1986)) has been codon optimized to improve translational efficiency of the HIV gagpol

proteins and reduce the risk of recombination between the transfer vector and HIV gagpol messenger RNA. Two hundred thirty-seven base pairs (237 bp) consisting of the gag pol overlap and cis-acting signals necessary for translation of pol (nucleotides 2583 to 2819 of SEQ ID NO: 12) were not optimized. The HIV gagpol sequence obtained using the codon optimization process does not differ at the amino acid level from the wildtype HIV gagpol sequence, but differs at the nucleotide level from the HIV gagpol sequence. A codon optimized HIV gag sequence is shown in Figures 8A-8E (pHDMHgpm2.seq) (SEQ ID NO:4). A codon optimized HIV pol sequence is shown in Figures 9A-9L (pHDMHgpm2.seq) (SEQ ID NO:10).

A plasmid comprising DNA sequences which encode codon optimized lentivirus gagpol proteins is also referred to herein as a packaging construct. This plasmid includes a promoter which drives the expression of the gagpol proteins, such as the human cytomegalovirus (hCMV) immediate early promoter. This plasmid is defective for the production of the viral envelope and accessory proteins tat, vif, vpr, vpu, nef and rev and the Rev response element (RRE). The packaging construct also does not contain viral sequences which are transcribed into mRNA, such as constitutive transport elements (CTEs).

10

A packaging construct comprising a codon optimized HIV gagpol is depicted in Figure 1 and in Figure 11. Figures 10A-10D depict the DNA sequence (SEQ ID NO:12) for the packaging construct pHDMHgpm2. This packaging construct (pHDMHgpm2) was constructed as follows: Plasmid pMDA.HIVgp mam was generated by chemical synthesis and PCR assembly (which is described in, for example, Stemmer et al., Gene, 164:49-53 (1995)) of 215 different oligonucleotides. The DNA sequence for pMDA.HIVgp mam is the same as the DNA sequence for pMDA.HIVgp jtg except for 4.3 kb which was codon optimized using the DNAStar program (LaserGene, Madison, WI). Two hundred thirty-seven base pairs (237 bp) consisting of the gag pol overlap and cis-acting signals necessary for translation of pol (nucleotides 2583 to 2819 of SEQ ID NO: 12) were not optimized due to dual reading

-13-

frame constraints. A NsiI site 5' of IN was preserved to aid fusion with wildtype sequences. Several single or double base pair silent mutations were introduced either to prevent potential splice donors and acceptors, or by the synthesis process. pMDA.HIVgp jtg was derived from HIV-1 strain NL4-3. The protease mutation that is present in the NL4-3 NIH GenBank sequence was then repaired (Figure 9B), changing the nucleotide present at position 2948 of SEQ ID NO:12 from a "G" to a "C", thereby producing the codon present at nucleotide positions 2948 to 2950 of SEQ ID NO:12 which encodes an arginine instead of the glycine present in the NL4-3 GenBank amino acid sequence. The resulting plasmid was named pMDHgpmam. The EcoRI-HindIII fragment of pMDHgpmam was inserted into pHDM2b, a high copy version of the pMD vector (Ory, D. et al., Proc. Natl. Acad. Sci. USA, 93(21):11400-11406 (1996)), to produce plasmid pHDMHgpm. The sequencing mutation that is present in the RNase domain of the NL4-3 NIH GenBank sequence was repaired (Figure 9H), changing the codon present at nucleotide positions 4724 to 4726 of SEQ ID NO:12 from "GGG" to "AAG", thereby producing a codon encoding a lysine instead of the glycine present in the NL4-3 GenBank amino acid sequence. The resulting plasmid was named pHDMHgpm2. Codon usage frequencies in the codon optimized gagpol open reading frame of the packaging construct pHDMHgpm2 are shown in Figure 2.

10

20

25

As used herein, a heterologous envelope protein permits pseudotyping of particles generated by the packaging construct and includes the G glycoprotein of vesicular stomatitis virus (VSV G) and the amphotropic envelope of the Moloney leukemia virus (MLV). A plasmid comprising a DNA sequence which encodes a heterologous envelope protein is also referred to herein as an envelope coding plasmid.

The terms "mammal" and "mammalian", as used herein, refer to any vertebrate animal, including monotremes, marsupials and placental, that suckle their young and either give birth to living young (eutharian or placental mammals) or are egg-laying (metatharian or nonplacental mammals). Examples of mammalian species include

humans and other primates (e.g., monkeys, chimpanzees), rodents (e.g., rats, mice, guinea pigs) and ruminents (e.g., cows, pigs, horses).

Examples of mammalian cells include human (such as HeLa cells, 293T cells, NIH 3T3 cells), bovine, ovine, porcine, murine (such as embryonic stem cells), rabbit and monkey (such as COS1 cells) cells. The cell may be a non-dividing cell (including hepatocytes, myofibers, hematopoietic stem cells, neurons) or a dividing cell. The cell may be an embryonic cell, bone marrow stem cell or other progenitor cell. Where the cell is a somatic cell, the cell can be, for example, an epithelial cell, fibroblast, smooth muscle cell, blood cell (including a hematopoietic cell, red blood cell, T-cell, B-cell, etc.), tumor cell, cardiac muscle cell, macrophage, dendritic cell, neuronal cell (e.g., a glial cell or astrocyte), or pathogen-infected cell (e.g., those infected by bacteria, viruses, virusoids, parasites, or prions).

Typically, cells isolated from a specific tissue (such as epithelium, fibroblast or hematopoietic cells) are categorized as a "cell-type." The cells can be obtained commercially or from a depository or obtained directly from an animal, such as by biopsy. Alternatively, the cell need not be isolated at all from the animal where, for example, it is desirable to deliver the virus to the animal in gene therapy.

15

To produce the cell lines of the present invention for producing a viral accessory protein independent lentivirus-derived retroviral vector particles, mammalian host cells are co-transfected with (a) a first plasmid comprising DNA sequence which encode lentivirus gagpol proteins, wherein said DNA sequence has been codon optimized by mutagenisis, as described above, to improve expression of the lentivirus gagpol proteins; and (2) a second plasmid comprising a DNA sequence which encodes a heterologous envelope protein, or a retroviral nucleotide sequence which comprises a DNA sequence of interest and lentivirus cis-acting sequences required for packaging, reverse transcription and integration, or both, under conditions appropriate for transfection of the cells.

20

25

In a particular embodiment, to produce the cell lines of the present invention for producing viral accessory protein independent HIV-derived retroviral vector particles mammalian host cells were cotransfected with (a) a first plasmid comprising DNA sequence which encode HIV *gagpol* proteins, wherein said DNA sequence has been codon optimized by mutagenisis, as described above, to improve expression of the HIV gagpol proteins; and (2) a second plasmid comprising a DNA sequence which encodes a heterologous envelope protein, or a retroviral nucleotide sequence which comprises a DNA sequence of interest and HIV cis-acting sequences required for packaging, reverse transcription and integration, or both, under conditions appropriate for transfection of the cells.

Virus stocks consisting of viral accessory protein independent lentivirus-derived, particularly HIV-derived, retroviral vector particles of the present invention are produced by maintaining the transfected cells under conditions suitable for virus production (e.g., in an appropriate growth media and for an appropriate period of time). Such conditions, which are not critical to the invention, are generally known in the art. See, e.g., Sambrook et al., Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor University Press, New York (1989); Ausubel et al., Current Protocols in Molecular Biology, John Wiley & Sons, New York (1998); U.S. Patent No. 5,449,614; and U.S. Patent No. 5,460,959, the teachings of which are incorporated herein by reference.

To generate viral accessory protein independent lentivirus-derived retroviral vector particles, mammalian host cells can be co-transfected with (a) a first plasmid comprising DNA sequence which encode lentivirus *gagpol* proteins, wherein said DNA sequence has been codon optimized by mutagenisis, as described above, to improve expression of the lentivirus gagpol proteins; (b) a second plasmid comprising a DNA sequence which encodes a heterologous envelope protein; and (c) a third plasmid comprising a DNA sequence of interest and lentivirus cis-acting sequences required for packaging, reverse transcription and integration. Alternatively, mammalian cells are

transfected with a plasmid comprising a codon optimized DNA sequence encoding a lentivirus gag protein and a plasmid comprising a codon optimized DNA sequence encoding a lentivirus pol protein, in place of the first plasmid comprising a codon optimized DNA sequence encoding both lentivirus gagpol proteins. Alternatively, mammalian host cells are transfected with a plasmid comprising a codon optimized DNA sequence encoding a lentivirus gag protein and a plasmid comprising a codon optimized DNA sequence encoding a lentivirus pol protein, in place of the first plasmid comprising a codon optimized DNA sequence encoding both lentivirus gagpol proteins.

In a particular embodiment, the invention relates to methods of producing viral accessory protein independent HIV-derived retroviral vector particles, comprising co-10 transfecting mammalian host cells with (a) a first plasmid comprising DNA sequence which encode HIV gagpol proteins, wherein said DNA sequence has been codon optimized by mutagenisis, as described above, to improve expression of the HIV gagpol proteins; (b) a second plasmid containing a DNA sequence which encodes a heterologous envelope protein; and (c) a third plasmid comprising a DNA sequence of 15 interest and HIV cis-acting sequences required for packaging, reverse transcription and integration. Alternatively, mammalian host cells are transfected with a plasmid comprising a codon optimized DNA sequence encoding a HIV gag protein and a plasmid comprising a codon optimized DNA sequence encoding a HIV pol protein, in place of the first plasmid comprising a codon optimized DNA sequence encoding both 20 HIV gagpol proteins.

Virus particles produced by the methods described herein, using a codon optimized HIV packaging construct produced as described herein, were compared by Western analysis with virus particles produced as described in Naldini *et al.*, *Science*, 272:263-267 (1996), using the packaging construct plasmid pCMVΔR8.2. Both the immunological reactivity and the proteolytic processing were confirmed to be indistinguishable.

-17-

5

10

2.4

20

25

A plasmid comprising a DNA sequence of interest and HIV cis-acting sequences required for packaging, reverse transcription and integration is also referred to herein as a transfer vector. A transfer vector, as used herein, refers to a vehicle which is used to introduce a DNA of interest into a eurkaryotic cell, particularly a mammalian cell.

Figure 3 depicts an example of a transfer vector.

DNA sequence of interest, as used herein, include all or a portion of a gene or genes encoding a nucleic acid product whose expression in a cell or a mammal is desired. In a particular embodiment, the nucleic acid product is a heterologous therapeutic protein. Examples of therapeutic proteins include antigens or immunogens, such as a polyvalent vaccine, cytokines, tumor necrosis factor, interferons, interleukins, adenosine deaminase, insulin, T-cell receptors, soluble CD4, growth factors, such as epidermal growth factor, human growth factor, insulin-like growth factors, fibroblast growth factors), blood factors, such as Factor VIII, Factor IX, cytochrome b, glucocerebrosidase, ApoE, ApoC, ApoAI, the LDL receptor, negative selection markers or "suicide proteins", such as thymidine kinase (including the HSV, CMV, VZV TK), anti-angiogenic factors, Fc receptors, plasminogen activators, such as t-PA, u-PA and streptokinase, dopamine, MHC, tumor suppressor genes such as p53 and Rb, monoclonal antibodies or antigen binding fragments thereof, drug resistance genes, ion channels, such as a calcium channel or a potassium channel, adrenergic receptors, hormones (including growth hormones) and anti-cancer agents. In another embodiment, the nucleic acid product is a gene product to be expressed in a cell or a mammal and which product is otherwise defective or absent in the cell or mammal. For example, the nucleic acid product can be a functional gene(s) which is defective or absent in the cell or mammal.

DNA sequence of interest includes DNA sequences (control sequences) which are necessary to drive the expression of the gene or genes. The control sequences are operably linked to the gene. The term "operably linked", as used herein, is defined to mean that the gene is linked to control sequences in a manner which allows expression

20

25

of the gene (or the nucleic acid sequence). Generally, operably linked means contiguous.

Control sequences include a transcriptional promoter, an optional operator sequence to control transcription, a sequence encoding suitable mRNA ribosomal binding sites and sequences which control termination of transcription and translation. In a particular embodiment, a recombinant gene encoding a desired nucleic acid product can be placed under the regulatory control of a promoter which can be induced or repressed, thereby offering a greater degree of control with respect to the level of the product produced.

As used herein, the term "promoter" refers to a sequence of DNA, usually upstream (5') of the coding region of a structural gene, which controls the expression of the coding region by providing recognition and binding sites for RNA polymerase and other factors which may be required for initiation of transcription. Suitable promoters are well known in the art. Exemplary promoters include the SV40, CMV and human elongation factor (EFI) promoters. Other suitable promoters are readily available in the art (see, e.g., Ausubel et al., Current Protocols in Molecular Biology, John Wiley & Sons, Inc., New York (1998); Sambrook et al., Molecular Cloning: A Laboratory Manual, 2nd edition, Cold Spring Harbor University Press, New York (1989); and U.S. Patent No. 5,681,735).

A DNA sequence of interest can be isolated from nature, modified from native sequences or manufactured *de novo*, as described in, for example, Ausubel *et al.*, *Current Protocols in Molecular Biology*, John Wiley & Sons, New York (1998); and Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*, 2nd edition, Cold Spring Harbor University Press, New York. (1989). DNA sequences can be isolated and fused together by methods known in the art, such as exploiting and manufacturing compatible cloning or restriction sites.

The packaging cell lines and viral particles of the present invention can be used, in vitro, in vivo and ex vivo, to introduce DNA of interest into a eukaryotic cell (e.g., a

-19-

mammalian cell) or a mammal (e.g., a human or other mammal or vertebrate). The cells can be obtained commercially or from a depository or obtained directly from a mammal, such as by biopsy. The cells can be obtained from a mammal to whom they will be returned or from another/different mammal of the same or different species. For example, using the packaging cell lines or viral particles of the present invention, DNA of interest can be introduced into nonhuman cells, such as pig cells, which are then introduced into a human. Alternatively, the cell need not be isolated from the mammal where, for example, it is desirable to deliver vial particles of the present invention to the mammal in gene therapy.

Ex vivo therapy has been described, for example, in Kasid et al., Proc. Natl. Acad. Sci. USA, 87:473 (1990); Rosenberg et al., N. Engl. J. Med., 323:570 (1990); Williams et al., Nature, 310:476 (1984); Dick et al., Cell, 42:71 (1985); Keller et al., Nature, 318:149 (1985); and Anderson et al., United States Patent No. 5,399,346.

10

15

20

Methods for administering (introducing) viral particles directly to a mammal are generally known to those practiced in the art. For example, modes of administration include parenteral, injection, mucosal, systemic, implant, intraperitoneal, oral, intradermal, transdermal (e.g., in slow release polymers), intramuscular, intravenous including infusion and/or bolus injection, subcutaneous, topical, epidural, etc. Viral particles of the present invention can, preferably, be administered in a pharmaceutically acceptable carrier, such as saline, sterile water, Ringer's solution, and isotonic sodium chloride solution.

The dosage of a viral particle of the present invention administered to a mammal, including frequency of administration, will vary depending upon a variety of factors, including mode and route of administration; size, age, sex, health, body weight and diet of the recipient mammal; nature and extent of symptoms of the disease or disorder being treated; kind of concurrent treatment, frequency of treatment, and the effect desired.

The teachings of all the articles, patents, patent applications and GenBank sequences cited herein are incorporated by reference in their entirety.

While this invention has been particularly shown and described with references to preferred embodiments thereof, it will be understood by those skilled in the art that various changes in form and details may be made therein without departing from the spirit and scope of the invention as defined by the appended claims.

#### **CLAIMS**

### What is claimed is:

- 1. A packaging cell line for producing a viral accessory protein independent HIVderived retroviral vector particle comprising:
- 5 a) a mammalian cell;
  - b) a first retroviral nucleotide sequence in the cell which comprises a coding sequence for a HIV gagpol, wherein said coding sequence has been mutagenized to improve expression of the HIV gagpol proteins;
  - c) a second retroviral nucleotide sequence in the cell which comprises the coding sequence for a heterologous envelope protein; and
  - d) a third retroviral nucleotide sequence in the cell which comprises a DNA sequence of interest and HIV cis-acting sequences required for packaging, reverse transcription and integration.
- 2. A packaging cell line of Claim 1 wherein the heterologous envelope protein is the G glycoprotein of vesicular stomatitis virus (VSV G).
  - 3. A packaging cell line of Claim 1 wherein the heterologous envelope protein is the amphotropic envelope of the Moloney leukemia virus.
  - 4. A packaging cell line of Claim 1 wherein the DNA sequence of interest encodes a heterologous therapeutic protein.
- 20 5. A packaging cell line comprising:
  - a) a mammalian cell;

15

- b) a first retroviral nucleotide sequence in the cell which comprises a coding sequence for a HIV gagpol, wherein said coding sequence has been mutagenized to improve expression of the HIV gagpol proteins; and
- c) a second retroviral nucleotide sequence in the cell which comprises a

  DNA sequence of interest and HIV cis-acting sequences required for packaging, reverse transcription and integration.
- 6. A packaging cell line of Claim 5 wherein the DNA sequence of interest encodes a heterologous therapeutic protein.
- 7. A packaging cell line comprising:
- 10 a) a mammalian cell;
  - b) a first retroviral nucleotide sequence in the cell which comprises a coding sequence for a HIV gagpol, wherein said coding sequence has been mutagenized to improve expression of the HIV gagpol proteins; and
  - c) a second retroviral nucleotide sequence in the cell which comprises the coding sequence for a heterologous envelope protein.
  - 8. A method of producing a packaging cell line for producing a viral accessory protein independent HIV-derived retroviral vector particle, comprising cotransfecting mammalian host cells with:
    - a first plasmid comprising a DNA sequence which encodes HIV gagpol proteins, wherein said DNA sequence has been mutagenized to improve expression of the HIV gag and pol proteins;
    - b) a second plasmid comprising a DNA sequence which encodes a heterologous envelope protein; and

- c) a third plasmid comprising a DNA sequence of interest and HIV cisacting sequences required for packaging, reverse transcription and integration.
- 9. A method of Claim 8 wherein the heterologous envelope protein is the G
   5 glycoprotein of vesicular stomatitis virus (VSV G).
  - 10. A method of Claim 8 wherein the heterologous envelope protein is the amphotropic envelope of the Moloney leukemia virus.
  - 11. A method of Claim 8 wherein the DNA sequence of interest is a heterologous therapeutic protein.
- 10 12. A method of producing a viral accessory protein independent HIV-derived retroviral vector particle comprising co-transfecting mammalian host cells with:
  - a) a first plasmid comprising a DNA sequence which encodes HIV gagpol proteins, wherein said DNA sequence has been mutagenized to improve expression of the HIV gagpol proteins;
- b) a second plasmid comprising a DNA sequence which encodes a heterologous envelope protein; and
  - a third plasmid comprising a DNA sequence of interest and HIV cisacting sequences required for packaging, reverse transcription and integration.
- 20 13. A method of Claim 12 wherein the heterologous envelope protein is the G glycoprotein of vesicular stomatitis virus (VSV G).

- 14. A method of Claim 12 wherein the heterologous envelope protein is the amphotropic envelope of the Moloney leukemia virus.
- 15. A method of Claim 12 wherein the DNA sequence of interest encodes a heterologous therapeutic protein.
- 5 16. A packaging cell line for producing a viral accessory protein independent lentivirus-derived retroviral vector particle comprising:
  - a) a mammalian cell;
  - b) a first retroviral nucleotide sequence in the cell which comprises a coding sequence for a lentivirus *gagpol*, wherein said coding sequence has been mutagenized to improve expression of the lentivirus gagpol proteins;
  - a second retroviral nucleotide sequence in the cell which comprises the coding sequence for a heterologous envelope protein; and
  - d) a third retroviral nucleotide sequence in the cell which comprises a DNA sequence of interest and lentivirus cis-acting sequences required for packaging, reverse transcription and integration.
  - 17. A packaging cell line of Claim 16 wherein the heterologous envelope protein is the G glycoprotein of vesicular stomatitis virus (VSV G).
- 18. A packaging cell line of Claim 16 wherein the heterologous envelope protein is the amphotropic envelope of the Moloney leukemia virus.
  - 19. A packaging cell line of Claim 16 wherein the DNA sequence of interest encodes a heterologous therapeutic protein.

PCT/US99/20675

- 20. A packaging cell line comprising:
  - a mammalian cell; a)

5

15

- b) a first retroviral nucleotide sequence in the cell which comprises a coding sequence for lentivirus gagpol, wherein said coding sequence-has been mutagenized to improve expression of the lentivirus gagpol proteins; and
- a second retroviral nucleotide sequence in the cell which comprises a c) DNA sequence of interest and lentivirus cis-acting sequences required for packaging, reverse transcription and integration.
- 21. A packaging cell line of Claim 20 wherein the DNA sequence of interest 10 encodes a heterologous therapeutic protein.
  - 22. A packaging cell line comprising:
    - a mammalian cell; a)
    - a first retroviral nucleotide sequence in the cell which comprises a b) coding sequence for lentivirus gagpol, wherein said coding sequence has been mutagenized to improve expression of the lentivirus gagpol proteins; and
    - c) a second retroviral nucleotide sequence in the cell which comprises the coding sequence for a heterologous envelope protein.
- A method of producing a packaging cell line for producing a viral accessory 20 23. protein independent lentivirus-derived retroviral vector particle, comprising cotransfecting mammalian host cells with:
  - a first plasmid comprising a DNA sequence which encodes lentivirus a) gagpol proteins, wherein said DNA sequence has been mutagenized to improve expression of the lentivirus gag and pol proteins;

- b) a second plasmid comprising a DNA sequence which encodes a heterologous envelope protein; and
- a third plasmid comprising a DNA sequence of interest and lentivirus
   cis-acting sequences required for packaging, reverse transcription and integration.
- 24. A method of Claim 23 wherein the heterologous envelope protein is the G glycoprotein of vesicular stomatitis virus (VSV G).
- 25. A method of Claim 23 wherein the heterologous envelope protein is the amphotropic envelope of the Moloney leukemia virus.
- 10 26. A method of Claim 23 wherein the DNA sequence of interest is a heterologous therapeutic protein.
  - 27. A method of producing a viral accessory protein independent lentivirus-derived retroviral vector particle comprising co-transfecting mammalian host cells with:
    - a) a first plasmid comprising a DNA sequence which encodes lentivirus gagpol proteins, wherein said DNA sequence has been mutagenized to improve expression of the lentivirus gagpol proteins;
    - b) a second plasmid comprising a DNA sequence which encodes a heterologous envelope protein; and
- a third plasmid comprising a DNA sequence of interest and lentivirus cis-acting sequences required for packaging, reverse transcription and integration.
  - 28. A method of Claim 27 wherein the heterologous envelope protein is the G glycoprotein of vesicular stomatitis virus (VSV G).

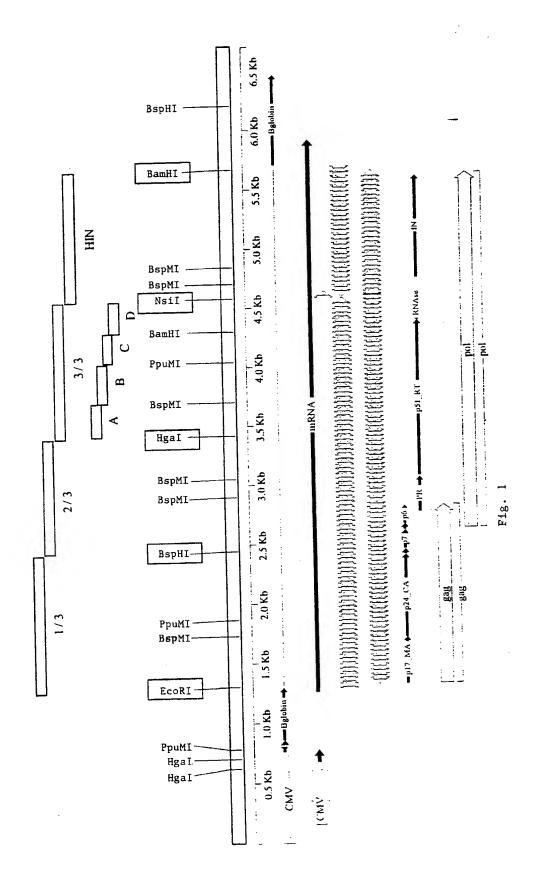
- 29. A method of Claim 27 wherein the heterologous envelope protein is the amphotropic envelope of the Moloney leukemia virus.
- 30. A method of Claim 27 wherein the DNA sequence of interest encodes a heterologous therapeutic protein.
- 5 31. A viral accessory protein independent HIV-derived retroviral vector particle produced by the method comprising co-transfecting mammalian host cells with:
  - a) a first plasmid comprising a DNA sequence which encodes HIV gagpol proteins, wherein said DNA sequence has been mutagenized to improve expression of the HIV gagpol proteins;
- b) a second plasmid comprising a DNA sequence which encodes a heterologous envelope protein; and
  - c) a third plasmid comprising a DNA sequence of interest and HIV cisacting sequences required for packaging, reverse transcription and integration.
- 15 32. A method of Claim 31 wherein the heterologous envelope protein is the G glycoprotein of vesicular stomatitis virus (VSV G).
  - 33. A method of Claim 31 wherein the heterologous envelope protein is the amphotropic envelope of the Moloney leukemia virus.
- 34. A method of Claim 31 wherein the DNA sequence of interest encodes a
   20 heterologous therapeutic protein.

- 35. A viral accessory protein independent lentivirus-derived retroviral vector particle produced by the method comprising co-transfecting mammalian host cells with:
  - a) a first plasmid comprising a DNA sequence which encodes lentivirus—
    gagpol proteins, wherein said DNA sequence has been mutagenized to
    improve expression of the lentivirus gagpol proteins;
    - b) a second plasmid comprising a DNA sequence which encodes a heterologous envelope protein; and
- a third plasmid comprising a DNA sequence of interest and lentivirus cis-acting sequences required for packaging, reverse transcription and integration.
  - 36. A method of Claim 35 wherein the heterologous envelope protein is the G glycoprotein of vesicular stomatitis virus (VSV G).
- 37. A method of Claim 35 wherein the heterologous envelope protein is the amphotropic envelope of the Moloney leukemia virus.
  - 38. A method of Claim 35 wherein the DNA sequence of interest encodes a heterologous therapeutic protein.
  - 39. Isolated DNA encoding a codon optimized HIV gagpol.
  - 40. Isolated DNA encoding a codon optimized HIV gag.
- 20 41. Isolated DNA of Claim 40 comprising the nucleotide sequence of SEQ ID NO:4.
  - 42. Isolated DNA encoding a codon optimized HIV pol.

-29-

- Isolated DNA of Claim 42 comprising the nucleotide sequence of SEQ ID NO:10.
- 44. A method of introducing a DNA sequence of interest into a mammal comprising introducing into said mammal a viral accessory protein independent HIV-derived retroviral vector particle comprising the DNA sequence of interest.
- 45. The method of Claim 44 wherein the mammal is a human.

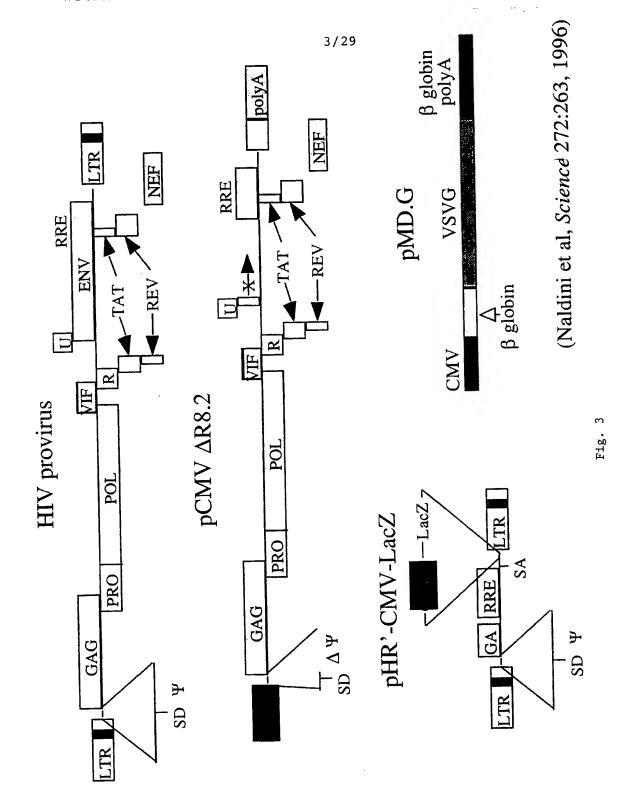
- 46. The method of Claim 44 wherein the DNA sequence of interest encodes a heterologous therapeutic protein.
- 47. A method of introducing a DNA sequence of interest into a mammal comprising the steps of:
  - introducing into cells a viral accessory protein independent HIV-derived
     retroviral vector particle comprising the DNA sequence of interest; and
  - b) returning the cells obtained in step a) to the mammal.
  - 48. The method of Claim 47 wherein the mammal is a human.
- 15 49. The method of Claim 47 wherein the DNA sequence of interest is a heterologous therapeutic protein.



Codon Usage Frequencies

100	III III	141	48		10	2 / 2	45.	2 <b>√</b>	ი (	78	6	13		<u>-</u>	57	٧.		1 5	3		74	56	\ <u>\</u>	25	3 4	9
nNI 4-3	pagnol	53	17		27	72	67	07 %	07 -	`	4	9		52	18			67	30		70	74	58	<u> </u>	2 4	01
Amino Acid		cca Pro(P)	ccc Pro(P)	cce Pro(P)	cen Pro(P)	300 Ser (S)	agu Ser (S)	usa Ser (S)			ucg Ser (S)	ucu Ser (S)		aca inr(I)	acc Thr (T)	acg Thr (T)	acu Thr (T)		(w)d11 88n	£ .	uac Iyr(Y)	uau Tyr (Y)	gua Val (V)		gue Val (V)	, a
mam		14	50	24	12	79	2.	i 		5	77	<u>8</u>	2	ი ;	26	58	٧.	, ,	ن ب	10	0 0	8.7	100		08	}
pNL4-3	gagpol	55	12	27	9	24	92	1		57	17	26	51	2 9	0		=	40	13	69	5 6	31	100		40	
Amino Acid		gga Gly(G)	ggc Gly(G)	ggg Gly(G)	ggu Gly(G)	cac Hi s(H)	cau Hi s(H)		11 (2)	ana IIe(I)	auc Ile(I)	auu Ile(I)	Cita [en(I)	one Lou(L)	כווכ דכוו(ד)	cng Leu(L)	cnn Leu(L)	una Leu(L)	ung Leu(L)	aga Lvs (K)		adg Lys (N)	ang Met (M)		uuc Phe (F)	į (
mam		13	53	17	17	10	18	9	37	21	; г	,	78	, (	7,	?	25	89	32		12	8	<del></del>		25	36
pNL4-3	gagpol	58	23	5	14	63	30	4	0	"	· C	>	27	73	0,	04	09	14	26		56	44			70	30
Amino Acid		gca Ala(A)	gcc Ala(A)	gcg Ala(A)	gcu Ala(A)	aga Arg(R)	agg Arg(R)	cga Arg(R)	cgc Arg(R)	cgg Arg(R)	con Ara(B)	cgu mg(m)	aac Asn(N)	aan Asn(N)	(A) V 000	gac Asp(D)	gan Asp(D)	uge Cys (C)	ugu Cys (C)		caa Gln(Q)	cag Gln(Q)	<del></del>		gaa Glu(E)	(1)(1)

Fig. 2



4/29

# Rev

- Regulates HIV gene expression by promoting cytoplasmic levels of unspliced and singly spliced mRNAs
- Postulated to affect splicing, stability, transport, and translation

Fig. 4

# Codon Optimization of HIV gagpol

- Remove A-rich instability elements
- Improve translational efficiency
- Reduce risk of recombination with transfer vector

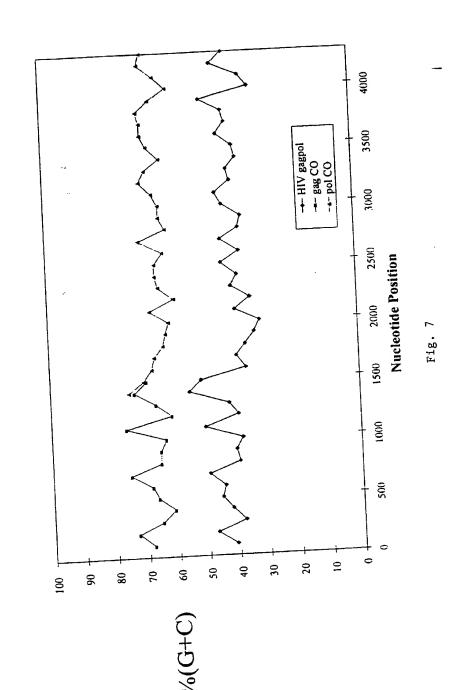
Fig. 5

# Inactivation of Inhibitory Sequences in gag

tta gac aag ata gag gaa gag caa aac aaa agt aag aaa aaa gca cag caa gca gca gct atg ggt gcg aga gcg tca gta tta agc ggg gga gaa tta gat cga tgg gaa aaa att cgg tta agg cca ggg gga aag aaa tat aaa tta aaa cat ata gta tgg gca agc agg gag <u>aca gta gca acc ctc tat tgt gtg cat caa agg ata gag ata aaa gac acc aag gaa gct</u> cta gaa cga ttc gca gtt aat cct ggc ctg tta gaa aca tca gaa ggc tgt aga caa ata ctg gga cag cta caa cca tc<u>c ctt cag aca gga tca gaa gaa ctt aga tca tta tat aat</u> <u>ں</u> ن U M<sub>2</sub> G Schwartz, S., et al. gac aca gga cac agc aat cag gtc agc caa aat tac C CC

Hg. 0

Nucleotide Content of HIV gagpoi



8/29

Alignment Report of Codon optimization (gag).MEG, using Clustal method with PAM250 residue weight table.

							810									
792	M	G	A	R	A	S	v	L	S	G	G	Е	L	D	К	NL4-3 genbank.SEO
792 1319	ATG	GGI			GCG											_
1319		G GGC	A. הכרכ	R . cec	A GCC	S	V GTG	L	S		G	E	L	D	K	pHDMHgpm2.seq
					. 300	100	313	CIG	100	. 660	. 666	. GAG	CTG	GAC	AAG	
		840										870				-
837	W	E	К	Ī	R	L	R	P	G	G	K	K	Q	Y	- 17	- NT 4 2 1
837	TGG	GAA	. AAA	ATT	CGG							AAA	CAA	TAT	AAA	NL4-3 genbank.SEQ
1364	W	E	K	I	R	L	R	₽	G	G	K	K	0	Y	к	pHDMHgpm2.seq
1364	TGG	GAG	AAG	ATC	CGC	CTG	CGC	CCC	GGC	GGC	AAG	AAG	CAG	TAC	AAG	. ,,
							900		<u>.</u>							-
882	L	К	H	I	v	W	 A	s	R	Е	L	E	R			
882					GTA					GAG	CTA	. GAA	CGA	F TTC	A GCA	NL4-3 genbank.SEQ
1409	L	K	H	Ι	V	W	А	S	R	E	L	E	R	F	A	pHDMHgpm2.seq
1409	CTG	AAG	CAC	ATC	GTG	TGG	GCC	TCC	CGC	GAG	CTG	GAG	CGC	TTC	GCC	[gp
		930										7 960				
927	v	N	Đ	G	L	L	E	Т	s	E	G	C	R			
927	GTT	AAT			CTT							TGT	AGA	Q CAA	ATA	NL4-3 genbank.SEQ
1454	v	N	₽	G	L	L	E	T	S	E	G	С	R	0	т	pHDMHgpm2.seq
1454	GTG	AAC	CCC	GGC	CTG	CTG	GAG	ACC	TCC	GAG	GGC	TGC	CGC	CAG	ATC	
							990					·				
972	L	G	Q	L	Q	P	s	L	Q	Т	G	S	E	Ē	L	NL4-3 genbank.SEQ
972	CTG	GGA	CAG	CTA	CAA	CCA							GAA	GAA	CTT	What's genbank.seg
1499	L	G	Q	L	Q	P	S	L	Q	T	G	s	Ξ	E	L	pHDMHgpm2.seq
1499	CTG	GGC	CAG	CTG	CAG	ccc	TCC	CTG	CAA	ACC	GGC	TCC	GAG	GAG	CTG	
	1	020							•		1	050				
1017	R	s	L	Y	N	T	Ī	A	v	L	Y		v	Н	Q	NL4-3 genbank.SEQ
1017	AGA	TCA	TTA	TAT	AAT	ACA	ATA						GTG			mua 3 dempany. 350
1544	R	S	L	Y	N	T	I	Α	V	L	Y	С	v	H	Q	pHDMHgpm2.seq
1544	CGC	TCC	CTG	TAC	AAC	ACC	ATC	GCC	GTG	CTG	TAC	TGC	GTG	CAC	CAG	
						10	080									
1062	R	I	D	v	К	D	T	К	E	A	L	D	ĸ	ī	E	NL4-3 genbank.SEQ
1062	AGG	ATA	GAT	GTA	AAA	GAC	ACC	AAG				GAT	AAG	ATA	GAG	J genbank.JEQ
1589	R	I	D	V	K	D	T	К	E	Α	L	D	K	I	E	pHDMHgpm2.seq
1589	CGC	ATC	GAC	GTG	AAG	GAC	ACC	AAG	GAG	GCC	CTG	GAC	AAG	ATC	GAG	
,	1:	110					<del></del> -				1	140				
1107	E	E	Q	N	К	s	К	К	К	Α	Q	Q	A		Δ	NI 1-2 manha-la gra
1107												CAA		GCA	A GCT	NL4-3 genbank.SEQ
1634	E	Ε	Q	N	K	S	K	K	K	Α	Q	Q	Α	Α	А	pHDMHgpm2.seq
1634	GAG	GAG	CAG	AAC	AAG '	rcc .	AAG .	AAG	AAG	GCC	CAG	CAG	GCC	GCC	GCC	- 244

Alignment Report of Codon optimization (gag).MEG, using Clustal method with PAM250 residue weight table.

							1	170	)										<del>.</del>	ura a manhank SEO
52 52 G 79 79 G	D GAC D GAC	T ACA T ACC	G G G GG	A AA 1 C AA	AC A	N AC N AC	S AGC S TCC	Q CA Q CA	G G	V TC / V TG /	S AGC S TCC	Q CAA Q CAG	N AA N AA	T T	Y AC Y AC	P CCT P CCC	I TA I TA	A G	TG V	NL4-3 genbank.SEQ
_		1												12	30					
97 -		200 N				G	Q	N	1	v	Н	Q	7	\ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \	I	S		P .	R	NL4-3 genbank.SEQ
97 ( 24 24	Q CAG Q CAG																			pHDMHgpm2.seq
-								126	50											
242 242	T ACT	L TT	A A	N AT (	A CA A	W TGG W	V GT V	A A	K AA K	V GTA V	V GT/ V	E A GA E	A G	E AG E	K AAG K	A GC A	T T	F TC F	S AGC S	NL4-3 genbank.SEQ pHDMHgpm2.seq
169 169	ACC	CT	G A	AC (	GCC	TGG	GT	G A	AG	GTG	GTO	G GA	.G G	AG	AAG		<u> </u>	. 10		
287		129 F		v	1	P	N	1	F	S	A	ן מיני	ra '	s	320 E	A GO		A GCC	T ACC	
287 814 814	CCA P	G G/ C G/	LA G	V TC	ATA I ATC	CCC	i A	(G ) (G )	F PTC	S TCC	A GC	c C	L IG '	s rcc	E GA	( G G	GC ·	A GCC	ACC T ACC	pHDMHgpm2.seq
									50											Turne SEC
332 332 859 859	P CC. P CC	A C	Q AA ( Q AG (	D GAT D GAC	L TTA L CTG	N AA N AA	T A	T CC T CC	M ATG M ATG	L CTA	N A AA N N	C A	T .CA T .CC	V GTG V GTG	G G G G	G G	G GA G GC	H CAT H CAC	Q CAA Q CAC	NL4-3 genbank.SEG pHDMHgpm2.seq G
		13													141	0				
1377 1377 1904 1904	7 GC	 \ :A @	A CC	M ATG M ATG	Q CAI Q CAI	A AI	G I	L TA L CTG	K AAA K AAG	E GA E GA	G A	r cc / r cc /	I ATC I ATC	N AA' N AA	r GA	E AG ( E AG (	E AA E AG	A GCT A GCC	GC:	NL4-3 genbank.SE A pHDMHgpm2.seq C
	_							1	440											
1422 1422 1949	2 1 2 Gi 9 G	E AA ' E AG '	W TGG W TGG	D GAT D GAC	R AG R	A T	L TG L TG	H CAT H CAC	CC D	A GI	G C	H AT H	A GCA A GCC	GG GG	G C	P CT P CC	I ATT I ATC	GC A GC	A CC	NL4-3 genbank.SE A pHDMHgpm2.seq
		14	70				·						D		15	00 A	G	T		NL4-3 genbank.S
146	17 17 G	G GC	Q CAG	M AT	F G AG	A G	e Aa E	CCA	AG	G G	GA /	AGT	GAC	: A'	ra (	CA A	GG# G	AC T		TT P pHDMHgpm2.seq

WO 00/15819 PCT/US99/20675

10/29 Alignment Report of Codon optimization (gag).MEG, using Clustal method with PAM250 residue weight table.

	·			٠,		,-										
						1	530							-		
1512	s	Т	L	Q	E	Q	I	G	W	M	T	Н	N	P	P	NL4-3 genbank.SEQ
1512						_		GGA	TGG	ATG	ACA		AAT			-
2039	s	T	L	Q	E	Q	I	G	W	M	T	Н	N	P	P	pHDMHgpm2.seq
2039	TCC	ACC	CTG	CAA	GAG	CAG	ATC	GGC	TGG	ATG	ACC	CAC	AAC	CCC	CCC	
	1	<del>,</del> 560									1	<del>7</del> 590				
1557		P	v	G	Ē	Ī	Y	к	R	W	I	I	L	G	L	NL4-3 genbank.SEQ
1557					GAA											o ge
2084	I	P	V	G	E	I	Y	K	R	W	I	I	L	G	L	pHDMHgpm2.seq
2084	ATC	CCC	GTG	GGC	GAG	ATC	TAC	AAG	CGC	TGG	ATC	ATC	CTG	GGC	CTG	
						1	620		<u> </u>							
1602	N	К	I	v	R	M	Y	5	P	Т	s	I	L	D	Ī	NL4-3 genbank.SEG
1602			_		AGA											, , , , , , , , , , , , , , , , , , ,
2129	N	K	I	V	R	М	Y	S	P	T	S	I	L	D	I	pHDMHgpm2.seq
2129	AAC	AAG	ATC	GTG	CGC	ATG	TAC	. TCC	CCC	ACC	TCC	ATC	CTG	GAC	ATC	
	1	650		<u></u>							1	680				
1647	R	<del>_</del>	G	P	К	E	P	F	R	D	Y	v	D	R	F	NL4-3 genbank.SEG
	AGA						CCC		AGA	GAC	TAT	GTA	GAC	CGA	TTC	•
2174	R	Q	G	P	K	E	P	F	R	۵	Y	V	D	R		pHDMHgpm2.seq
2174	CGC	CAG	GGC	CCC	AAG	GAG	CCC	TTC	CGC	GAC	TAC	GTG	GAC	CGC	TTC	
			<u></u>			1	710								<del></del>	-
1692	Y	К	T	L	R	A	E	Q	А	S	Q	Е	v	К	. N	NL4-3 genbank.SE
1692	TAT	AAA			AGA											
2219	Y	K	T	L	R	A	Ë	Q	A	S	Q	E	V CTD	K	N	pHDMHgpm2.seq
2219	TAC	AAG	ACC	CTG	CGC	GCC	GAG	CAG	GCC	100	CAG	GAG	GIA	AAG	AAC	
	1	740			- "						1	770				
1737	W	М	T	Ε	Т	L	L	V	Q	N	A	N	P	D	Ç	
1737					ACC											
2264	W	M	T	E	T ACC	L	L	ν CTC	Q	N	A	N	P	D D	C TGC	pHDMHgpm2.seq
2264	TGG	ATG	ACC	GAG	ACC	CIG	CIG	GIG		Anc		rate				-
						]	800									_
1782	к	Т	I	L	к	A	L	G	Р	G	A	T	L	E	E	NL4-3 genbank.SE
	AAG															
2309	K AAG	T	I	L	K	A	L	G								pHDMHgpm2.seq
2309	AAG	ACC	ATC	CTG	AAG	GCC	CTG	GGC		GGC			CIG	GAG	GAG	
		830										.860				_
1827	M	М	T	A	С	Q	G	٧	G	G	P	G				NL4-3 genbank.SE
1827	ATG	ATG	ACA	GCA	TGT	CAG	GGA	GTG	GGG				CAT	AAA	GCA	•
2354	М	M	T	A	С	Q	G	v	G	G	P	G	H	К	А	pHDMHgpm2.seq
2354	ATG	ATG	ACC	GCC	TGC	CAG	GGC	GTG	GGC	GGC	CCC	GGC	CAC	AAG	GCC	

Fig. 8C

11/29

Alignment Report of Codon optimization (gag).MEG, using Clustal method with PAM250 residue weight table.

	·							_									
							18	90									
.872	R	v	L		A	E	A	M	S	Q	٧.	T	N	P	A	T	NL4-3 genbank.SEQ
.872	AGA	GTT	TTG	G	CT (	SAA (	GCA .	ATG .	AGC	CAA	GTA .	ACA	AAT	P	GCT A	ACC	pHDMHqpm2.seq
	~	17	T		Δ	F	А	M	S	Q	V	Т	1.4	-	^	r	bunudhum. 204
399	CGC	GTG	CTG	3 G	CC (	GAG	GCC	ATG	TCC	CAA	GIC	MCC	AAC	CCC	300	ACC	
-		1										1	950		•		
		20				**		N	F	R	N	Q	R	К	T.	v	NL4-3 genbank.SEQ
917	I ATA	M	I		Q	K	G	N አልጥ	ኒኒኒኒኒ E	AGG							•
	-		т		^	K	G	N	F	R	N	Q	ĸ	L.	1	٧	pHDMHgpm2.seq
444	I ATC	ATG	AT(	c. (	ZAG	AAG	GGC	AAC	TTC	CGC	AAC	CAG	CGC	AAG	ACC	GTG	
444	——	<u></u>															-
							1	980									
962	ĸ	С	F		N	С	G	K	E	G	H	I	A	K	N		NL4-3 genbank.SEQ
962	AAG	TGT	TT	C Z	AAT		GGC	AAA	GAA	GGG	CAC	ATA	A	K	W	C	pHDMHgpm2.seq
489	ĸ	С	F	•	N	C	G	K	E	G	H	I					
489	AAG	TGC	TT	C	AAC	TGC	GGC	AAG	GAG		CAC		<del></del>				_
		010										2	040				_
		A	P	,	R	К	К	G	С	W	К	С	G	к	E		NL4-3 genbank.SE(
2007	R AGG	GCC	. cc	T.	AGG	AAA	AAG	GGC	TGT	TGG	AAA	TGT	GGA	AAG	GAA	GGA	
	_			3	כ	ĸ	ĸ	G	С	W	K	C	G	v	Ľ.	G	pupunghus . 264
2534	CGC	GCC	: cc	C	CGC	AAG	AAG	GGC	TGC	TGG	AAG	TGC	GGC	AAG	GAG	GGC	:
								2070									-
				<u>.</u>	75	D		T	Ē	R	Q		N	E	L	G	NL4-3 genbank.SE
2052	H	Q	יו דיריי	n rc	K	GAT	TGT	ACT	GAG	AGA	CAG		AAT	TTT	TTA	GGG	;
2052		_			v	ח	С	T	Ε	R	Q	А	7.4	2		J	5.mmap.a.
2579 2579	CAC	CAC	G A2	rG	AAA	GAT	TGT	ACT	GAG	AGA	CAG	GCT	AA?	TTT	TTA	. GGG	3
2313																	_
	;	2100	•										2130				 NL4-3 genbank.SE
2097	К	I	7	W	P	S	H	K	G	R	. p	G	N	ያ መመጣ መመመ	r cm.	Q r car	
2097	AAG	AT	C T	GG					GG	A AGO	نمان و P	G G	N N	TT? F	L L	Q	pHDMHqpm2.seq
2624	K	I	1	W	Р	S	H	K - 226	G - ~~;	R							•
2624	AAG	AT	C T	GG	CCT	TCC	CAL		, GG								
								21,60						<del></del>			
2142		R		P	E	P	Т	A	P	P	Ē	E					NL4-3 genbank.Si
2142	: S : AG		_		~~		A AC	A GC	c cc	A CC	A GA	A GA	G AG	C TT	C AG	G TT	T
		-		_		ס	- 1	A	Ρ.				-	-	• • •	-	P. 11-1-1-20
2669	AG	C AG	A C	CA	GAG	CC	A AC	A GC	C CC	A CC	A GA	A GA	G AG	C TT	C AG	G TT	ı
		27.0											2220	)			<del> </del>
		219									- 7			D	Т	ת	NL4-3 genbank.S
2187	7 G	E	:	Ε	T	T	T	P T CC	S - ידיר	Q دے ت	ת מנים	د ر <u>ه</u> ح	ے جی ج	G CC	G AT	A GA	iC
	7 GG	_		~	e e	T T	יוף			U	r.		_		_	-	
2714	4 G 4 GG	_ E	;	<u>ت</u> 200	1	ת. המיב	ז זר י	ጥ ሮሮ	כ דר	y CA	 G AA	G CA	G GA	G CC	G AI	A GA	vc
271	4 GG	G GA	VY G	λG	المكالم	- AC	a ac							_			

12/29

Alignment Report of Codon optimization (gag).MEG, using Clustal method with PAM250 residue weight table.

						2	250			<del></del>						
2232 2232 2759 2759	K	E	L	Y	P	L TTA L TTA	А	S	L	R	S	L	F	G	9	NL4-3 genbank.SEQ
		280	<del></del> -		<del></del> _		•				•					·
2277 2277			S TCG	S TCA	Q CAA	TAA	•									NL4-3 genbank.SEQ
2804 2804	D GAC	P CCC	S TCG	5 TCA	Q CAA	TAA										pHDMHgpm2.seq

13/29

Alignment Report of Codon Optimization (pol) MEG, using Clustal method with PAM250 residue weight table.

	T										<del> v -</del>								· · · · · · · · · · · · · · · · · · ·
	20	90											212						wr 4 2 manhank SEO
2087 -	F	E	R	Е		D	L	Α	F	Р.	Q			K	A	R			NL4-3 genbank.SEQ
2087	r TTT 1	TT	AGG					CC '	TTC	CCA	CAA	. G(	36 A 3	K	A	R	<b>G G</b>	E.	pNL4-3.seq
2085	F	F	R	Ε		D .	L TG (	A	E	CC.V	Q CAA								<b>P.1.</b>
2085 '	TTT 1							A	F	P	Q		G	K	Α	R		Ε	pHDMHgpm2.seq
2612	F TTT 1	F	R	E	, ,	D Domotion	L	-A	ው ተ	CCA	CAA				GCC	AG	G G	AΑ	
2612	TTT	LLL	AGG	GΑ	A G	WI (													
-								50	_										on a grant SEO
2132	F	s	s	E	Ξ.	Q	Т	R	A	N	S		P :	T	R	P AG		E iag	NL4-3 genbank.SEQ
2132	F TTT	TCT						AGA	GCC	AAC	AG(		P P	T	R	F	 }	E	pNL4-3.seq
2130	F TTT	S	S	E	: 	Q	T	R	A	N AAC	AGG								
							ACC T	AGA R	A	N	s		P	T'	R	F	₹	Ε	pHDMHgpm2.seq
2657	F TTT	S ~~~	S TrCI	י נים ג	5 26 (	Q TAG	ACC	AGA	GCC	AAC	AG	c 0	cc i	ACC	AGA	A	3A (	AG	
2657			107																
•		180											22	210					6 1 650
2177	L	<del>'</del> Q	v	1	W	G	R	D	N	N	5		r	S 	E		A.		NL4-3 genbank.SEQ
2177	L CTT	CAG	GT'	r T	GG (	GGA	AGA	GAC	AAC	AAC	TC:	C	CTC	TCA	GAP		⊶. A	G	pNL4-3.seq
2175	_	_	7.7	,	r.7	_	ס	n	N	IN.	J			-	-			-	
2175	L CTT	CAG	GT				AGA	GAC	AAC	AA(	S TO		L	S	E		A.	G	pHDMHgpm2.seq
2702	L CTT	Q	v		W	G	R	D	N	N ימה י	است. د	· - (	CTC						
2702	CTT	CAG	GT	тт	GG	GUA	AGA	anc	Purc										•
							2	240											
2222		D	R		Q	G	T	v	S	F	5	;	F	2	Q		I	T	
2222	GCC	GAS	r AG	A C	AA.	GGA	ACT	GTA	TC	TT	T A	GC.	TTC	CCT	CA	g A	TC	ACT	-VIA 3 coc
2220		_			$\sim$		·	v	- 5	r		,	<u></u>	-	¥		_	_	
2220	GCC	GA:	r AG	A	ZΑA	GGA	ACT	GTA	TC	TT	T A	GC .	TTC	CCI	· CA	دين	T	AC1	pHDMHgpm2.seq
			_		_	_	177	17	9	}*			2	-	×		-	-	
2747 2747	GCC	GA'	r Ac	A C	CAA	GGA	ACT	GTA	A TC	C TT	T A	jĊ	TTC	Ç ( )	. CA	G F	120	AC I	· 
		2270	·										2	300	-				-
					B	Б	L	v	Т			K	I	G	G		Q	L	
2267	L	W 	ر در	2	R	בכר	CTC	GT	c ac	A AI	'A A	AG	ATA	GG	G GG	G (	CAA	TTA	<del>f</del>
2267				_	~	•	Ť	W	T			n.		•	_	•	×	_	
2265	L	W TG	יה כי	AG (	CGA	ccc	CTC	GT	C AC	A A	ra a	AG	ATA	GG	G GG	G	CAA	TT?	A
2265 2792				_	-		•	17	'11'			Λ.	1	0	•	•	×.	_	F
2792	. הלהלו ה	TG	G C	AG .	CGA	CCC	CTC	GT	C AC	A A.	ca a	ΑG	ATC	GG'	T GO	C (	CAG	CT(	3
2132																			
								2330									7		NL4-3 genbank.SEQ
2312	K	E	1 .	A	L	L	D	T			Α -	D	D	T	, c:	v r:5	ترىنىت ت	G.P.	7
2312	K RAA	G GA	A G	CT	CTA	TT	GA:	. AC	A GO	A G	CA G	ΑT	GAT	AU m	A G	 . /	1 1 M	F.	pNL4-3.seq
2310	) K	G G#			CTA	TT!	A GA	r AC	A GC	aA G •	CA G A	D	D	 Τ	, ,	v .	L	E	pHDMHgpm2.seq
	7 AA																		
2837	7 AA	G G#	∖G G	CC	CTG	CT	∍ GA	L AC	ر ال	<i>-</i>	(		··			-			

14/29

Alignment Report of Codon Optimization (pol).MEG, using Clustal method with PAM250 residue weight table.

			_														, 00,01	e weight table.
		23	60					•	<del></del>				·····	390		*		
235	57 -	——. Е	M	N	L	р								390				
235		- Aaa					G				ζ :	₽ 	K	M	I	0	5	S NL4-3 genbank.SI
235			M	N	L	P	G	n AG R	A To	G A	AA C	CA A			AT.	A GO	G G	GA .
235	55 G/						رود	ת מנו	ו א יייר	G AA	\ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \	P 	K	M	I	G	; (	pNL4-3.seg
288	32 E	2	M	N	L	P	G	r Au R	W IG	O AA	VA CC	JA, A				A GG	G G	SA
288	12 G/	AG A	TG A		CTG	ccc	GGG		י ייינ	ו א בע א	יר כנ יר כנ	: :	K	M	I	G	(	pHDMHgpm2.seq
							-0.		0 10	IG AA	iG CC		MG	ATG	ATO	GG	C G	SC
								2420	)									Thinks.
240	2 1		G G	G	F	I	К											
240	2 AT	T G						V	 	Q	Y		D	Q	I	L	I	NL4-3 genbank.SE
240				G .	F	I	K	V. GE	שט א	A CA	G TA	T G	AT		ATA	CT	C AI	'A
240	0 AT	T G	_							Q N Ch	Y	- -	D 	Q	I	L	I	pNL4-3.seq
292	7 r	(	3							A CÃ								A -
292	7 AT	C GO	GC G	GC 1	TC	ATC	AAA	CT (	רכי	Q 	Y TD	<b>.</b> .	D	Q	Ι	L	1	pHDMHgpm2.seq
						• • • •	100	. 610	- (6	C CA	G IA	.C G.	AC	CAG	ATC	CT	G AT	c
		245	0								<u>-</u>		24	180				<del>-</del>
244	7 E	I	. (	<del></del> -	G	Н	К	A	I	G	T							<del></del>
2447	7 GA.	A AI					AAA	GCT	ידע י	A GG1	ים בי	ים ב	√ ר תים	L	V	G	P	NL4-3 genbank.SE
2445	5 E	I	. (	3	G	H	к	A	I	. UG.	T		/ /	L				r
2445	GA/	TA A	C TO	SC G	GA (	CAT	AAA	GCT	' ATA	A GG1	r AC	י מים	י מי	רידית ה	Cm v	G	P	pNL4-3.seq
2972		_	,	-	G	п	ĸ	A	т	G	T	7	<i>t</i>	T		_		
2972	GA(	3 AT	C TO	SC G	GC (	CAC	AAG	GCC	ATC	GGC	: AC	. G1	' 'G (	TG.	GTG.	G	P	pHDMHgpm2.seq
																		-
2492	т							510										-
	-	P			N 	I	I	G	R	N	L	L	,	T	Q	I	G	- NL4-3 genbank.SEÇ
2490	ACA T	P	r Gr	CA	AC A	ATA .	ATT	GGA	AGA	AAT	CTC	TT	G A	CT	CAG	ATT	GGC	;
2490	-	-	٧		ν,	Τ.	1	G	R	N	L	L	ı	T	0	т	G	DMT 4 2
3017	T	P	v		N N	I.		GGA	AGA	AAT				CT	CAG	ATT	GGC	
	_					.m.⊂ `:	I	G	R	N	L	L	_	T	Q	I	G	pHDMHgpm2.seq
	ACC		, 01	3 74	nc A	110	AIC	GGC	CGC	AAC	CTG	CT	G A	.cc	CAG	ATC	GGC	
	2	2540							<del></del>				25					-
2537		T	-		<u> </u>								25				_	_
2537		ACT	L. יידטידטי	N מממ		E mm (	P	I	S	P	I	Ε		T	V	P	v	- NL4-3 genbank.SEQ
2535	C	T	L	A.A.A N		TT (	DUC.	ATT	AGT	CCT			G A	CT (	GTA	CCA	GTA	3 DEQ
2535		ACT	_	N AA		F TT (	P	I	S	P	I	E		T	v	P	V	pNL4-3.seq
3062	c	T	L	י הת א		r (	P.		AGT	CCT	ATT					CCA	GTA	•
3062						ייר כ	.c.c	I	S	P	I	E		Г	V	P	V	pHDMHgpm2.seq
_			•••			10 0		AIC	100	ccc	ATC	GAC	3 A(	CC (	TG (	CCC	GTG	•
							26	00									<del></del> .	
2582	ĸ	L	К	Б		<del></del>	M	D			77	<del></del>						
2582							። ጥር ረ	ייי מב	G	P	K	V	ŀ	ζ .	Q	W	P	NL4-3 genbank.SEQ
2580	К	L	K	P			M.	D	G	ULA D	MAA						CCA	
2580 3107			AAG	CC	A GG	א ב	ጉር <i>ር</i>	ነው ነ	G GGC	CCD P	K	C TP TP	4		Q	W	P	pNL4-3.seq
3107	K	L	К	P	. G	12 }	M	D.	G	P	MAA.	GIT						
3107				ccr	. GG	C A	TG (	ב. באר י	عدر ع	כככ ב	K nnn	CTIC	K		Q N	W	P	pHDMHgpm2.seq
					_ 00	A		arc (	300	الل	MAA	GTC	AA	G C	AG T	'GG	CCC	•

15/29

Alignment Report of Codon Optimization (pol).MEG, using Clustal method with PAM250 residue weight table.

		<del></del>							<del></del>			<del></del>				
	2	630									2	660				
2627	L	T	Ε	Ε	К	I	К	A	L	V	E	I	C	T	E	NL4-3 genbarrk.SEQ
2627	TTG	ACA	GAA	GAA	AAA	ATA	AAA	GCA	TTA	GTA			TGT	ACA	GAA	-
2625	L	T	E	E	K	I	К	A	L	V	Ε	I	C	T	E	pNL4-3.seq
2625					AAA											
3152	L	T	E	E	K	I	K	A	L	V	E	I	C	T	E	pHDMHgpm2.seq
3152	CTG	ACC	GAG	GAG	AAG	ATC	AAG	GCC	CTG	GIG	GAG	ATC	TGC	ACC	GAG	
2670							690		72							ur i a seste i ara
2672 2672	M ATG	E	K	E	G	K	I Ture	5 TC2	K AAA		GGG	P CCT	E GAA	N AAT		NL4-3 genbank.SEQ
2670	M	E	K	E	GGA	K	I	S	K	I	G	P	E	N	5	pNL4-3.seg
2670				_					AAA							Firms aloned
3197	M	E	K	E	G	K	I	S	К	I	G	P	E	N	P	pHDMHqpm2.seq
3197									AAG	ATC	GGC					
												т				
	2	720									2	750				
2717	Y	N	T	P	v	F	A	I	К	К	К	۵	s	Т	К	NL4-3 genbank.SEQ
2717	TAC	AAT	ACT	CCA	GTA	TTT	GCC	ATA	AAG	AAA	AAA	GAC	AGT	ACT	AAA	
2715	Y	N	T	P	V	F	A	I	K	K	К	D	S	Т	K	pNL4-3.seq
2715	TAC	AAT			GTA				AAG							.0
3242	Y	N	T	P	V	F	A	I	K	K	K	D	S	T	K	pHDMHgpm2.seq
3242	TAC	AAC	ACC	CCC	GTG	TTC	GCC	ATC	AAG	AAG	AAG	GAC	TCC	ACC	AAG	
						2	780								·	
27.52			7.7	<del>-</del>					E	L	N	К	R	T		NI 1-2 combank SEO
2762 2762	W	R	K nnn	L	V GTA	D GAT	ωω-C E	R							CZ2 Q	NL4-3 genbank.SEQ
2760	W	R	K	L	V	D	F	R	E	L	N	К	R	T	Q.	pNL4-3.seq
2760			AAA	_	GTA				GAA						-	p
3287	M	R	K	L	v	D	F	R	Ξ	L	N	K	R	T	Q	pHDMHqpm2.seq
3297					GTG			CGC	GAG	CTG	AAC	AAG	CGC	ACC		
		-	<del></del>							<del></del>		<del></del> -				
	2	810									2	940				
2807	D	F	W	E	7/	Q	L	G	I	5	H	P	А	G	L	NL4-3 genbank.SEQ
2807	GAT	TTC	TGG	GAA	GTT	CAA	TTA	GGA	ATA	CCA	CAT	CCT	GCA	GGG	TTA	
2805	D	F	W	E	٧	Q	L	G	I	P	H	P	Α	G	L	pNL4-3.seq
2905	GAT	TTC	TGG	GAA	GTT	CAA			ATA							
3332	D	F	M	E	V	Q	L	G	I	5	H	P	A	G	L	pHDMHgpm2.seq
3332	GAC	TTC	TGG	GAG	GTG	CAG	CTG	GGC	ATC	ccc	CAC	ccc	GCC	GGC	CTG	
			,				870									
2052				7.5				37		D	17	G	D		Y	NL4-3 gembank.SEQ
2852 2852	K	Q	K	K	5 TC3	CT2			L CTG							ME4-2 deupaux.256
			AAA K	AAA K	S	V	T	V	L	D	V	G	D	A	Ă.	pNL4-3.seq
2850 2850	מבמ	Q CAG	A A A													0.00q
	K	Q	K	K	S	V	T	v	L	D	V	G	D	A		pHDMHqpm2.seq
3377	AAG	CAG														

16/29

Alignment Report of Codon Optimization (pol).MEG, using Clustal method with PAM250 residue weight table.

		<del></del>														
		2900										2930				
2897	-	S	V	P	L	D	K		F	R	К	Y	T	A	F	- NL4-3 genbank.SEQ
2897	TT	TC	A GT	r cc	TTP	GA?	C AA	A GA	C TTO	AGG	AAG	TAT	r Aci	r GCZ	ኒ ፓጥፕ	, what's dempark.sed
2895	) F	S	V	P	L	D	K	D	F	R	K	Y	Т	A	F	pNL4-3.seq
2895			A GT		TTA	GA7	LAA 1	A GAG	C TTC	AGG	AAG	TAT	CAC 1	GC/	TTT	, F2. 3.3cq
3422		S	V	P	L	D	K	D	F	R	K	Y	T	А	F	nHDMHanm2
3422	TTC	TCC	GT	G CCC	CTG	GAC	: AA	GA(	C TTC	CGC	: AAG	TAC	ACC	GCC	TTC	;
																-
							2960									,
2942	_	I	P	S	I	N	N	E	T	P	G	I	R	Y	Q	NL4-3 genbank.SEQ
2942				AG1		. AAC	: AA	GA	ACA	CCA	GGG	ATI	' AGA	TAT	CAG	
2940	_	I	P	. S	I	N	N	E	T	P	G	I	R	Y	Q	pNL4-3.seq
2940 3467				AGI		AAC			ACA			ATI	AGA	TAT	CAG	•
	-	I	. ccc	S	I I	N	N	E	T	P	G	I	R	Y	Q	pHDMHgpm2.seq
3407	ACC	AIC		. 100	, AIC	AAC	AAC	GAG	ACC	ccc	GGC	ATC	CGC	TAC	CAG	
		2990			·············			· ·			3	3020			·	-
2987	Y	N	v	L	P	Q	G	W	K	G	s	P	A			-
2987	TAC	AAT	GTG	CTT	CCA							CCA	. הרא	I ATTA	F	NL4-3 genbank.SEQ
2985	Y	N	V	L	P	Q	G	W	К	G	S	P	A	I	F	mNT 4 2
2985	TAC	AAT	GTG	CTT	CCA	CAG	GGA	TGG						ATA	יי. ד	pNL4-3.seq
3512	Y	N	v	L	P	Q	G	W	K	G	S	P	А	r	च	pHDMHgpm2.seg
3512	TAC	AAC	GTG	CIG	CCC	CAG	GGC	TGG	AAG	GGC	TCC	CCC	GCC	ATC	TTC	phornigpmz.seq
		·					3050		<del></del>	<del></del>						•
3032	-0	С	S													
3032	_			M ATC	T ACA	K	I	L	E	P	F	R	K	Q	N	NL4-3 genbank.SEQ
3030	Q	C	S	M	T	K	I	L	GAG E				AAA			
3030	_				ACA				GAG	P	F TTT	R	K	Q	N	pNL4-3.seq
3557	Q	C	S	М	T	K	I	L	E E	P	F	AGA R	AAA K			- 1170.00
3557	CAG	TGC	TCC	ATG	ACC						TTC	CGC	AAG	Q CAG	N AAC	pHDMHgpm2.seq
	3	080									3	110				
3077	P	D	I	V	I	Y	Q	Y	М	D	D	L	Y	v	G	NL4-3 genbank.SEQ
3077					ATC	TAT	CAA	TAC	ATG	GAT	GAT	TTG				9 genbankible
3075	P	D.	I	V	I	Y	Q	Y	M	D	D	L	Y	V	G	pNL4-3.seq
3075		GAC			ATC				ATG		GAT	TTG	TAT	GTA	GGA	- 4
3602	P	D	I	V	I	Y	Q	Y	М	D	D	L	Y	ν	G	pHDMHgpm2.seq
3602	CCC	GAC	ATC	GIG	ATC	TAC	CAG	TAC	ATG	GAC	GAC	CTG	TAC	GTG	GGC	-
					<del></del>	3	140						<del></del> -			
3122	s	D	L	E	I.	G	Q	Н	R	T	К	ī	E-		<del></del> -	
3122							CAG	СУТ	AGA	ACA	ΑΔΑ	בדם בדם	E GAG	E	L	NL4-3 genbank.SEQ
3120	s	D	L	E	I	G	Q	Н	R	T	K	I	E E	GAA E		mN7 4 2 -
3120	TCT						CAG	CAT	AGA	ACA	AAA	ATA	GAG	GAA	L CTG	pNL4-3.seq
3647	S	D	L	E	I	G	Q	H	R	T	ĸ	I	E	E.	Ť.	pHDMHqpm2.seq
3647	TCC	GAC	CTG	GAG	ATC	GGC			CGC	ACC	AAG	ATC	GAG	GAG	CTG	Parangpaiz.seq

17/29

Alignment Report of Codon Optimization (pol).MEG, using Clustal method with PAM250 residue weight table.

		<del></del>										<del></del>	<del></del>	···		<del></del>
	3	170									3	200				
3167	R	Q	Н	L	L	R	W	G	F	T	T	P	D	K		NL4-3 genbank.SEQ
3167						AGG										
3165 3165	R	Q CAA	H	L	L	R AGG	W	G	thathate E	T	T	P	D	K	K	pNL4-3.seq
3692	R	Q	H	L	L	R	M	G	F	T	T	P	D	AAA K	AAA K	pHDMHqpm2.seq
3692																buntatabus raed
						3	230									
3212	H	Q	K	E	P	P	F	L	W	М	G	Y	E	L	Н	NL4-3 genbank.SEQ
3212	CAT	CAG	AAA			CCA			TGG		GGT	TAT	GAA	CTC	CAT	
3210	H	Q	K	E	P	P	F	L	W	M	G	Y	Ε	L	H	pNL4-3.seq
3210				GAA		CCA										. Uma est
3737 3737	H	CAG	K AAG	EGAG	CCC	P CCC	F TTC	L CTG	W TGG	M ATG	GGC	Y TAC	E	L	H	pHDMHgpm2.seq
3737	<b>W</b> 10	00	10.0	<b></b>	000	000	1.0	010	100		-	1110	G, 10	OIG	CAC	
	3	260									3	290				
3257	P	D	K	W	T	V	Q	P	I	v	L	P	E	K	D	NL4-3 genbank.SEQ
3257	CCT	GAT	AAA	TGG	ACA	GTA	CAG	CCT	ATA	GTG	CTG	CCA	GAA	AAG	GAC	, ,
3255	P	D	K	W	T	V	Q	P	I	V	L	P	E	K	D	pNL4-3.seq
3255						GTA										
3782 3782	P	D	K	W	T	V CTC	Q CAC	5 5	I	V CTC	L	P	E	K	D	pHDMHgpm2.seq
3102	CCC	GAC	AAG	100	ACC	GIG	CAG	CCC	AIC	GIG	CIG	cec	GAG	AAG	GAC	
				<del></del>		3	320									
3302	s	W	T	v	N	D	Ī	Q	К	L	v	G	К	L	N	NL4-3 genbank.SEQ
3302	AGC	TGG	ACT	GTC	AAT	GAC	ATA	CAG	AAA	TTA	GTG	GGA	AAA	TTG	AAT	, ,
3300	s	W	T	V	N	D	I	Q	K	L	V	G	K	L	N	pNL4-3.seq
	AGC					GAC										
3827	S	W	T	V	N	D	I	Q	K	L	V	G	K	L 	N	pHDMHg <b>p</b> m2.seq
3827	TCC	TGG	ACC	GTG	AAC	GAC	ATC	CAG	AAG	CrG	GTG	GGC	AAG	CTG	AAC	
	3	350									3	380				
3347	W	A		Q	r	_ <u></u>	Α	G	I	К	v	R	Q	L,	С	NIA-3 conbank SEO
3347		GCA				TAT										NL4-3 genbank.SEQ
3345	W	A	s	Q	I	Y	A	G	I	K	v	R	Q	Ľ	С	pNL4-3.seg
3345	TGG	GCA	AGT	CAG	ATT	TAT	GCA	GGG	ATT	AAA	GTA	AGG	CAA	TTA	TGT	•
3872	W	Α	S	Q	I	Y	Α	G	I	K	V	R	Q	L	С	pHDMHgpm2.seq
3872	TGG	GCC	TCC	CAG	ATC	TAC	GCC	GGC	ATC	AAA	GTC	CGC	CAG	CTG	TGC	
							<del></del>						···			
							410									
3392		-	L	R			К		L							NL4-3 genbank.SEQ
3392			-													
3390 3390		L CTT	Critic L	R AGG	G GGA		K aaa		L CTA	T			V GTD		L CTA	pNL4-3.seq
3917		L	L	R	G		K	A	L				V		L	pHDMHapm2.sea
3917														_		f zadkume i n n d

18/29

Alignment Report of Codon Optimization (pol).MEG, using Clustal method with PAM250 residue weight table.

		<del></del>			<del></del> -				·	····		<del></del>		<del></del>	<del></del>	
		1440										3470				_
3437	T	E	E	A	E	L	E	L	A	E	N	R	E	I	L	NL4-3 genbank.SEQ
3437 3435	ACA T	GAA E	GAA E	GCA												
3435				A GCA	E	L	E	L	A	E	N	R	E	I	L	pNL4-3.seq_
3962	T	E	E	A	E	L	E	L	A	GAA. E	AAC N	AGG R	GAG E		CTA	
	ACC													I ATC	L	pHDMHgpm2.seq
							<b></b>	010	-	w		000	OAG	AIC	CIG	
						3	500									-
3482	К	Ε	P	v	Н	G	v	Y	Y	D	P	s	К	D	L	NL4-3 genbank.SEQ
3482	AAA	GAA	CCG	GTA	CAT	GGA	GTG	TAT	TAT	GAC					TTA	a demodiff.DDA
3480	K	Ε	P	V	H	G	v	Y	Υ.	D	P	S	K	D	L	pNL4-3.seq
3480	AAA			GTA		GGA	GTG	TAT	TAT	GAC	CCA	TCA	AAA	GAC	TTA	•
4007	K	E	Б	V	H	G	٧	Y	Y	D	P	S	K	D	L	pHDMHgpm2.seq
4007	AAG	GAG	CCC	GTG	CAC	GGC	GTG	TAC	TAC	GAC	ccc	TCC	AAG	GAC	CTG	
	3	530					···	<del></del>				560				
3527		A	E		Q	К										
3527				ATA			Q CAG	G	Q CAA	G	Q	W	T	Y	Q	NL4-3 genbank.SEQ
3525	I	Α	E	I	Q	K	Q	G	Q	G	Q	W	T	Y	Q	DNI 4 2 and
3525	_			ATA												pNL4-3.seq
4052	I	Α	E	I	Q	K	Q	G	Q	G	Q	W	T	Y	Q	pHDMHapm2.sec
4052	ATC	GCC	GAG	ATC	CAG	AAG		GGC								pribitingphiz . seq
														<del></del>		
						3	590									
3572	I	Y	Q	E	P	F	K	. N	L	К	T	G	K	Y	A	NL4-3 genbank.SEQ
3572				GAG		TTT							AAA	TAT	GCA	
3570	I	Y	Q	E	P	F	К	N	L	ĸ	T	G	K	Y	A <sub>.</sub>	pNL4-3.seq
3570 4097	ATT			GAG												
4097		Y TAC	CAC.	E	CCC B	F	K	N	L	K	T	G	K	Y	A	pHDMHgpm2.seq
1037	AIC.	IAC	CAG	GAG	CCC	110	AAG	AAC	CIG	AAG	ACC	GGC	AAA	IAC	GCC	•
	3	620									3	650				
3617	R	M	К	G	A	Н	T	N	D	v	К	<u>-</u>	L	T	E	NL4-3 genbank.SEQ
3617	AGA												TTA			J gemank.JEQ
3615	R	M	K	G	Α	Н	T	N	D	v	К	Q	L	T	E	pNL4-3.seq
3615	AGA	ATG	AAG	GGT	GCC	CAC	ACT	AAT	GAT	GTG	AAA	CAA	TTA	ACA	GAG	•
4142	R	M	К	G	Α	H	Ţ	N	D	V	K	Q	L	T	E	pHDMHgpm2.seq
4142	CGC	ATG	AAG	GGC	GCC	CAC	ACC	AAC	GAC	GTG	AAG	CAG	CTG	ACC	GAG	
						3	680							-		
2662	<del></del>	<u> </u>		<del></del>	<del></del>											
3662	A	V	Q	K		A	T	E	S	I	V	I	W	G	K	NL4-3 genbank.SEQ
3662 3660		GTA V														
3660			CDD	K aaa	I ATA	A GCC	T	E GAA	S	I ata	V CTA	I ara	₩ TGG	G	K	pNL4-3.seq
4187	A	V	Q	K	I	A	T	E	S	I	V	I	M	GGA		pHDMHgpm2.seq
4187															AAG	hunudhus sed
- <del>-</del> -										•		•			. 74.0	

19/29

Alignment Report of Codon Optimization (pol).MEG, using Clustal method with PAM250 residue weight table.

•	•															
	37	710				-	<del></del>				37	40				
3707	T	P	К	F	К	L	P	I	-Q	К	E	T	W	E	A	NL4-3 genbank.SEQ
3707	ACT	CCT	AAA	TTT			CCC	ATA	CAA	AAG	GAA	ACA	TGG	GAA	GCA	
3705	T	P	К	F	K	L	P	I	Q	ĸ	E	T	W	E	A	pNL4-3.seq
3705	_			TTT	AAA	TTA	CCC	ATA	CAA	AAG	GAA	ACA	TGG	gaa	GCA	
4232	T	P	к	F	K	L	P	I	Q	K	E	T	W	E	Α	pHDMHgpm2.seq
4232	ACT	ccc	AAG	TTC	AAG	CTG	CCC	ATC	CAG	AAG	GAG	ACC	TGG	GAG	GCC	
						3	770	<del>-</del>								
3752	W	W	T	E	Y	W	<u> </u>	A	T	W	I	P	·E	W	E	NL4-3 genbank.SEQ
3752				GAG	TAT	TGG	CAA	GCC	ACC	TGG	ATT	CCT	GAG	TGG	GAG	
3750	W	W	T	E	Y	W	Q	Α	T	W	I	P	E	W	E	pNL4-3.seq
3750	TGG	TGG	ACA	GAG	TAT	TGG	CAA	GCC	ACC	TGG	ATT	CCT	GAG	TGG	GAG	
4277	W	W	Т	E	Y	W	Q	Α	T	W	I	P	E	W	E	pHDMHgpm2.seq
4277	TGG	TGG	ACC	GAG	TAC	TGG	CAG	GCC	ACC	TGG	ATC	CCC	GAG	TGG	GAG	
	3	800						· · · · · · · · · · · · · · · · · · ·			3	830				
2707			N	Т	P	P	L	v	К	L	W	Y	Q	L	E	NL4-3 genbank.SEQ
3797	F	CTPC	774	י אַרַר י	CCT										GAG	-
3797 3795	F	V	N	T	5	P	L	v	К	L	W	Y	Q	L	E	pNL4-3.seq
3795	ահանան Ե	CTC.	ייעמ	ACC	CCT					TTA	TGG	TAC	CAG	TTA	GAG	•
4322	F	V	N	T	P	Р	L	V	K	L	W	Y	Q	L	E	pHDMHgpm2.seq
4322	TTC	GTG	AAC	ACC	CCC	CCC	CTG	GTG	AAG	CTG	TGG	TAC	CAG	CTG	GAG	
							3860									-
				I	ı	G	A	E	Т	F	Y	v	D	G	A	NL4-3 genbank.SEQ
3842	K	E	CCC		ATA							GTA	GAT	GGG	GCA	
3842		E	P	I	I	G	A	E	T	F	Y	V	D	G	Α	pNL4-3.seq
3840	K				ATA	-				TTC	TAT	GTA	GAT	GGG	GCA	
3840 4367	K	E	P	I	I	G	A	E	Т	F	Y	V	D	G	Α	pHDMHgpm2.seq
4367	AAG	GAG	ccc	ATC	ATC	GGC	GCC	GAG	ACC	TTC	TAC	GTG	GAC	GGC	GCC	
												3920			·	<u>.</u>
		3890														- NIA-2 gonbank SEO
3887	A	N	R	Ε	T	K	L	G	K	A	G	Y	V ייברים	T	D D	NL4~3 genbank.SEQ
3887	GCC	AAT			A ACT						455 Z	L IAI	V V	T ACI	GAC D	pNL4-3.seq
3885	Α	N	R	E	Т	K	L	G	K	A	G	. Y			' GAC	
3885					A ACT					A GCA	G G	Y	. Gla	T AC	. GAC	pHDMHqpm2.seq
4412	A	N	R	Ε	T	K	L	G	K							
4412	GCC	AAC	CG	C GAG	G ACC	. AAG		ادادا و	, AAG						, one	, 
							3950									
3037	R	G	R	Q	К	v		P	L	T						
3932	AGA	GG	A AG	a cai	A AAA	A GTT	r GT	ccc	CTA	ACC	GAC	: AC	A AC	AA?	CAC	3
2020		G	P	0	К	V	V	P	L	T	Ð	T	T	N	Q	pnL4-3.seq
3930	AGA	. GG	A AG	A CA	A AA	GT7	r GT	cco	CTA	A ACC	GAC	ACA	A ACA	A AAC	CAC	3
4457	<b>D</b>	G	R	0	К	V	V	Р	L	Т	ט	Т	T	N	Q	phomingpinz.seq
4457	CGC	GG	CG	C CA	G AA	GTO	G GT	CC(	CTC	G AC	GAG	ACC	C ACC	CAA	CA(	3
,																

20/29

Alignment Report of Codon Optimization (pol).MEG, using Clustal method with PAM250 residue weight table.

		<del></del>						<del></del>				7	<del></del>			
		3980										4010				
3977		T	E	L	Q	A	I	H	L	A	L	Q	D	S		
3977					A CAA									TC	G GGA	
3975 3975		T	E	L	Q	A	I	Н	L	A	L	Q	D	S	G	pNL4-3.seq
4502		T	E	L	A CAA Q	A.	I AT	CAT H	L	A GC1		CAG			G GGA	
	AAG	-	_	_						A GCC	L Tro	. CV	D	S	G	pHDMHgpm2.seq
									. 010		. 010	CAL	. GAC	. 100	- 660	•
							4040									<del></del>
4022	L	E	v	N	I	v	T	D	s	Q	Y					
4022												A GCA	L Tro	G	I A ATC	NL4-3 genbank.SEQ
4020		Ε	V	N	I	v	T	D D	S	Q	Y	A	L	G G	I	pNL4-3.seq
4020	TTA	GAA	GTA	AAC	ATA	GTG	ACA	GAC	TCA						ATC	
4547		E	v	N	I	v	T	D	s	Q	Y	A	L	G	I	pHDMHapm2.sea
4547	CTG	GAG	GTG	AAC	ATC	GTG	ACC	GAC	TCC	CAG	TAT	GCA	TTG	GGC	ATC	
		1070														-
40												1100	<del></del> -			<u>.</u>
4067	-	Q	A	Q	P	D	К	S	E	S	E	L	V	S	Q	NL4-3 genbank.SEQ
4067 4065															CAA	
4065		Q CAA	GC <sub>2</sub>	Q CAA	P CCA	D Car	K	S ACT	E	S	E	L	V	S	Q	pNL4-3.seq
4592	I	Q	A	Q	P	D	K	S	E	S	E	L	A GIC	AG1		- Image - 2
	ATC													TCC	Q CAG	pHDMHgpm2.seq
				<del>.</del>			<del></del>									•
						4	130									
4112	I	I	E	Q	L	I	K	К	E	К	٧	Y	L	Α	W	NL4-3 genbank.SEQ
4112	ATA I	I										TAC			TGG	
	ATA		E	CAG	L TTD	I ATA	K	K	E GDD	K	∨ V	Y	L	A	W	pNL4-3.seq
4637	I	I	E	Q	L	I	K	K	E	K	V	Y	L			-110161 0
4637														A GCC	W TGG	pHDMHgpm2.seq
															100	
	4	160		•							4	190				
4157	v	Б	A	Н	К	G	I	G	G	N	E	Q	v	D	G	NL4-3 genbank.SEQ
4157	GTA	CCA	GCA	CAC	AAA	GGA	ATT	GGA	GGA						GGG	a deimanniona
4155	V	P	A	H	K	G	I	G	G	N	E	Q	v	D	K	pNL4-3.seq
4155		CCA				GGA			GGA			CAA	GTA	GAT	AAG	•
4682	V	P	A	H	K	G	I	G	G	N	E	Q	٧	D	K	pHDMHgpm2.seq
4682	GTG	ccc	GCC	CAC	AAG	GGC	ATC	GGC	GGC	AAC	GAG	CAG	GTG	GAC	AAG	
							220				·					
4202	т.	v	s	A	G		R	<i>V</i>	v							
4202													D Carr	GGN	Σων Τ	NL4-3 genbank.SEQ
4200		v	S	A	G	I	R	K	V	L	F	L	D	G	I	pNL4-3.seq
4200																Pinna 3.3ed
4727	L	V	S	A,	G	I	R	K	V	L	F	L	D	G	I	pHDMHqpm2.seg
4727	CTG	GTG	TCC	GCC	GGC .	ATC	CGC	AAG	GTG	CTG	TTC	CTG	GAC	GGC	ATC	

21/29

Alignment Report of Codon Optimization (pol).MEG, using Clustal method with PAM250 residue weight table.

		T										Т				
		250									42	280				
4247	D	К	A	Q	E	E	H	E	K	Y	Н	S	N	W	R	NL4-3 genbank.SEQ
4247	GAT	AAG	GCC		GAA											-NT 4 2
4245	D	K	A	Q	E	E	H	E	K	Y Tar	CAC	S NGT	N aat	W TGG	R	pNL4-3.seq
4245	GAT				gaa E	GAA E	H	E	K	Y	H	S	N	W	R	pHDMHgpm2.seq
4772	D	K	A	Q CNG	GAG											huntudhum
4772	GAC	AAG	GCC	CAG	GAG	CAC	<b>W10</b>	O. 10				•				
						4	310							* .		
4202				s	D	F	N	L	P	P	v	v	A	ĸ	E	NL4-3 genbank.SEQ
4292 4292	A GCA	M ATC			GAT											y
4290	A	M	A	S	D	F	N	L	P	P	v	٧	Α	К	E	pNL4-3.seq
4290					GAT			CTA	CCA	CCT	GTA	GTA	GCA	AAA	GAA	
4817	A	M	Α	S	D	F	N	L	P	P	V	V	A	K	E	pHDMHgpm2.seq
4817	GCC	ATG	GCC	TCC	GAC	TTC	AAC	CTG	CCC	CCC	GTG	GTG	GCC	AAG	GAG	
		-									<u></u> -	1				
	4	340									4	370				-
4337	I	V	A	S	С	D	K	С	Q	L	K	G	E	A	М	NL4-3 genbank.SEQ
4337	ATA	GTA	GCC	AGC	TGT	GAT	AAA	TGT								
4335	I	V	Α	, S	С	D	K	С	Q	L	K	G	E	A	M	pNL4-3.seq
4335					TGT								GAA E	A	ATG M	pHDMHqpm2.seq
4862 4862	I	V	A	S	C	D	K	C TCC	Q CAG	L CTG	K AAG	G GGC				hupundhus . sed
4862	ATC	GTG	GCC	100	160	GAC	AAG	160	CAG	0.0	15.0	-	00	-		
						4	400									•
4000				v	D	С	5	Р	G	r	W	Q	L	D	С	NL4-3 genbank.SEQ
4382 4382	H	G CGN	Q CAA	CTA	GAC											, <u>.</u>
4380	H	G	0	v	D	c	S	P	G	I	W	Q	L	D	С	pNL4-3.seq
4380	CAT	GGA	CAA		GAC	TGT	AGC	CCA	GGA	ATA	TGG	CAG	CTA	GAT	TGT	
4907	н	G	0	V	D	С	S	₽	G	I	W	Q	L	D	С	pHDMHgpm2.seq
4907	CAC	GGC	CAG	GTG	GAC	TGC	TCC	CCC	GGC	ATC	TGG	CAG	CTG	GAC	TGC	
												<del></del>				-
	4	430										460				_
4427	T	H	L	E	G	К	V	I	L	V	A	٧	н	V	A	NL4-3 genbank.SEQ
4427	ACA	CAT	TTA	GAA	GGA	AAA	GTT	ATC						GTA		
4425	T	Н	L	Ε	G	K	V	I	L	V	A	V	Н	V	A	pNL4-3.seq
4425	ACA	CAT			GGA											
4952	T	H	L	E	G	K	V	I	L	V	A.	V GTG	H CAC	V GTG	A . ccc	pHDMHgpm2.seq
4952	ACC	CAC	CTG	GAG	GGC	AAG	GTG	ATC	. CrG	GIG	ا با	. 616	, wat	. 516	GCC	
							1490									-
			<del></del>	<del></del>				17	I	P	A	E	T	G	Q	_ NL4-3 genbank.SEQ
4472	S AGT	G	Y To Tr	I מידהי	E	A CCA	E	V GTA								_
		وني G	Y.		E E	A.	E E	V V	I	P	A	E	T	. 555 G	Q	
4470	S AGT	י הכיז י	ኒ ኒ	ב מידמי	GAD.	GCA									_	
4997		G	Y IAI	I	E	A	E	v	Ī	P	A	Ε	T	G	Q	pHDMHgpm2.seq
4997	TCC	GGC	TAC	ATC	: GAG								ACC	GGC	CAG	
2001		. 550														

22/29

Alignment Report of Codon Optimization (pol).MEG, using Clustal method with PAM250 residue weight table.

		1520									4	550				
4517	E	т	A	Y	F	L	L	К	L	A	G	R	W	P	v	- NL4-3 genbank.SEQ
4517	GAA	ACA	GCA	TAC	TTC	CTC	TTA	AAA	TTA						GTA	"" a deimair. 250
4515	E	T	Α	Y	F	L	L	K	L	Α	G	R	W	P	v	pNL4-3.seq
4515	GAA	ACA	GCA	TAC	TTC	CTC	TTA	AAA	TTA	GCA	GGA	AGA	TGG	CCA	GTA	
5042	E	T	Α	Y	F	L	L	К	L	A	G	R	W	P	v	pHDMHgpm2.seq
5042	GAG	ACC	GCC	TAC	TTC	CTG	CTG	AAG	CTG	GCC	GGC	CGC	TGG	CCC	GTG	
										<del></del> ,						_
						4	1580									
4562	K	T	V	H	T	D	N	G	S	N	F	T	S	T	T	NL4-3 genbank.SEQ
4562	AAA	ACA	GTA	CAT	ACA	GAC	AAT	GGC	AGC	AAT	TTC	ACC	AGT	ACT	ACA	3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3
4560	K	T	V	H	T	D	N	G	S	N	F	T	S	T	T	pNL4-3.seq
4560					ACA	GAC	AAT	GGC	AGC	AAT	TTC	ACC	AGT	ACT	ACA	•
5087	K	T	V	Н	T	D	N	G	S	N	F	T	s	T	T	pHDMHqpm2.seq
5087	AAG	ACC	GTG	CAC	ACC	GAC	AAC	GGC	TCC	AAC	TTC	ACC	TCC	ACC	ACC	•
		610		-							<i>-</i>	640				
4600					<del></del>											
4607	V	K	A	A	C	W	W	A	G	I	K	Q	Ē	F	G	NL4-3 genbank.SEQ
4607		AAG			TGT		TGG								GGC	
4605	Cana V	K	A	A	C	W	W	A	G	I	K	Q	E	F	G	pNL4-3.seq
4605 5132	V	AAG K													GGC	
5132			A	A	C TGC	W TCC	W TCC	A	G	I	K	Q	E	F	G	pHDMHgpm2.seq
3132	919	AAG	GCC	GCC	190	166	166	GCC	GGC	AIC	AAG	CAG	GAG	TTC	GGC	
						4	670				<del></del>	·	-			
4652	I	P	Y	N	Р	Q	 S	Q	G	v	I	E	s	М		NT 4 - 2
4652					ccc	_										NL4-3 genbank.SEQ
4650	I	P	Y	N	P	Q	s	Q	G	v	I	E	s	М	N	pNL4-3.seq
4650	ATT	ccc	TAC	AAT	ccc											para-3.seq
5177	I	P	Y	N	P	Q	s	Q	G	V	I	E	s	М	N	pHDMHqpm2.seq
5177	ATC	CCC	TAC	AAC	ccc	CAG	TCC		GGC	GTG						piibi nigpiaz . seq
												<del>-</del>				
	4	700									4	730				
4697	K	E	L	К	K	I	I	G	Q	V	R	D	Q	A	E	NL4-3 genbank.SEQ
4697	AAA	GAA	TTA	AAG	AAA	ATT	ATA	GGA	CAG	GTA	AGA	GAT	CAG	GCT		,
4695	K	E	L	K	K	I	I	G	Q	V	R	D	Q	Α	E	pNL4-3.seq
4695		gaa.	TTA	AAG	AAA	ATT	ATA	GGA	CAG	GTA	AGA	GAT	CAG	GCT	GAA	•
5222	K	Ε	L	K	К	I	I	G	Q	Λ	R	D	Q	Α	E	pHDMHgpm2.seq
5222	AAG	GAG	CTG	AAG	AAG	ATC	ATC	GGC	CAA	GTC	CGC	GAC	CAG	GCC	GAG	
						4	760		<del></del>							
4740	<del></del>	<del>-</del> -					٠							<u> </u>		
4742	H	L	K	T	A	V		M				I		N	F	NL4-3 genbank.SEQ
4742					GCA											
4740	H	L	K	T	A	V	Q	M	A	V V	F	I	H	N	F	pNL4-3.seq
4740 5267																
	H	L	K	T	A	V	Q	M	A	CTC	F	I	H	N	F	pHDMHgpm2.seq
3201	CAC	CIG	MMG	ACC	GCC	GIG	CAG	AIG	GCC	GIG	TTC	ATC	CAC	AAC	TTC	

23/29

Alignment Report of Codon Optimization (pol).MEG, using Clustal method with PAM250 residue weight table.

		- <del></del>										<del>,</del>				<del></del>
	4	790									4	820				
4787	K	R	К	G	G	I	G	G	Y	\$	A	G	Ε	R	I	NL4-3 genbank.SEQ
4787				GGG									GAA			
4785	K	R	K	G	G	I	G	G	Y	S	A	G	E	R	I	pNL4-3.seq
4785				GGG G	GGG	ATT		GGG			GCA A	GGG	gaa E	AGA R		
5312 5312	K AAG	R	K	_	_	_	G	_	Υ	S TCC					I ATC	pHDMHgpm2.seq
3312	ANO		rato	444	300	AL C	300	000			000	-	0,,0			
						4	850		-							•
4832		D	I	Ī		T	D	ī	Q	T	К	E	L	Q	ĸ	NL4-3 genbank.SEQ
4832				ATA												
4830	V	D	I	r	A	T	D	I	Q	T	K	E	L	Q	K	pNL4-3.seq
4830	GTA	GAC	ATA	ATA	GCA	ACA	GAC	ATA	CAA	ACT	AAA	GAA	TTA	CAA	AAA	
5357	V	D	I	I	Α	T	D	I	Q	T	K	E	L	Q	K	pHDMHgpm2.seq
5357	GTG	GAC	ATC	ATC	GCC	ACC	GAC	ATC	CAG	ACC	AAG	GAG	CTG	CAG	AAG	
		880									4	910				
4077		1											<u> </u>			NT ( 2 )- GEO
4877 4877	Q	I	T	K AAA	I	Q CAA	N	f TTT	R	V GTT	Y TAT	TAC	R AGG	D D	S	NL4-3 genbank.SEQ
4875	Q	I	T	K	I	Q	N	F	R	V	Y	Y	R	D	S	pNL4-3.seg
4875	_		_	AAA			AAT	TTT		GTT	TAT		AGG			part sibed
5402	Q	I	T	К	I	Q	N	F	R	٧	Y	Y	R	D	S	pHDMHgpm2.seq
5402	CAG	ATC	ACC	AAG	ATC	CAG	AAC	TTC	CGC	GTG	TAC	TAC	CGC	GAC	TCC	
						4	940									
4922	R	D	P	V	W	K	G	P	Α	K	L	L	W	K	G	NL4-3 genbank.SEQ
4922				GTT												nNT 4 2 and
4920	R	D CAT	P	Cutur	W TCC	K AAA	G GGA	P	A CCA	K	L CTC	L CTC	W TGG	K AAA	G GGT	pNL4-3.seq
4920 5447	AGA R	D	CCA P	V	M	K	G	P	A	K	L	L	W	K	G	pHDMHqpm2.seq
5447		_		GTG												piisi nigpiia i seq
	4	970									5	000				
4967	Ε	G	A	٧	v	I	Q	D	N	s	D	I	K	V	V	NL4-3 genbank.SEQ
4967	GAA	GGG	GCA	GTA	GTA		CAA	GAT	AAT				AAA			
4965	E	G	A	V	V	I	Q	D	N	S	D	I	K	V	V	pNL4-3.seq
4965				GTA												-11D)(11
5492	E	G	A	V CMC	V	I	Q	D	N	S	D CAC	I	K	V CTC	V GTC	pHDMHgpm2.seq
5492	GAG	GGC	GCC	GTG	GIG	AIC	CAG	GAC	AAC	100	GAC	VI.C	ANG	313	GIG	
						5	030									•
5012	P	R	R	К	A	К	I	I	R	D	Y	G	К	Q.	М	NL4-3 genbank.SEQ
5012																o gemann.oug
		R	R	К	A	К	I	I	R	Đ	Y	G	к	Q	М	pNL4-3.seq
5010	₽															-
5010				AAA	GCA	AAG	ATC	ATC	AGG	GAT	TAT	GGA	AAA	CAG	ATG	
	CCA P	AGA R	AGA R		Α	K	I	I	R	D	Y	G	К	Q	M	pHDMHgpm2.seq

24/29

Alignment Report of Codon Optimization (pol).MEG, using Clustal method with PAM250 residue weight table.

		060					-					-			
		060									5	090			
5057	Α	G	D	D	С	v	A	S	R	Q	D	E	D	4	NL4-3 genbank.SEQ
5057	GCA	GGT	GAT	GAT	TGT	GTG	GCA	AGT	AGA	CAG	GAT	GAG	GAT	TAA	MEN-3 Gemank.seQ
5055	A	G	D	D	С	V	A	S	R	Q	D	E	D		pNL4-3.seq
5055	GCA	GGT	GAT	GAT	TGT	GTG	GCA	AGT	AGA	CAG	GAT	GAG	GAT	TAA	pana-5.seq
5582	Α	G	D	D	С	v	A	S	R	Q	D	E	ם	Francisco .	pHDMHgpm2.seq
5582	GCC	GGC	GAC	GAC	TGC	GTG	GCC	TCC	CGC	CAG	GAC	GAG	GAC	TAA	phorngpitz.seq

Fig. 9L

				_		
AGCTTGGCCC	ATTGCATACG	TTGTATCCAT	ATCATAATAT	GTACATTTAT .	ATTGGCTCAT	60 120
OPPOST A COOPE	አሮሮሮሮሮ፯ሞርሞ	ጥርልሮልጥፕGAጥ '	<b>TATTGACTAG</b>	TTATTAATAG	TAATCAATIA	120
-caamaxam	አ ርጥጥር አጥልርር	CCATATATGG :	AGTTCCGCGT	TACATAACTT	ACGGIAAAIG	180 240
accessore.	CTCACCCCCC	AACGACCCCC	GCCCATTGAC	GTCAATAATG	ACGIAIGIIC	300
	CCCNATACCC	V CALALAC A JALL	GACGTCAATG	GGTGGAGTAT	TIACGGIAAA	360-
COCCOO COO	CCCACTACAT	CAAGTGTATC	ATATGCCAAG	TACGCCCCCT	AT LOWCRICK	420
- mas access 3	A TO CO	TCCCATTATG	CCCAGTACAT	GACCTTATGG	GACTITICCIA	480
~~~~~~~~~~~	<b>ለ አጥርጥ አ ሮርጥ አ</b>	TTAGTCATCG	CTATTACCAT	GGTGATGCGG	TITIGGCAGI	540
2020022000	CCCTCCATAG	CGGTTTGACT	CACGGGGATT	TCCAAGTCTC	CACCCCATIG	600
ACGTCAATGG	GAGTTTGTTT	TGGCACCAAA	ATCAACGGGA	CTTTCCAAAA	TGTCGTAACA	660
- 000000000	አመመር እሮሮሮ ል ል	ATGGGCGGTA	GGCGTGTACG	GTGGGAGGTC	IMIMIAMGEN	720
	አ ሮመሮ አ አ ሮሮሮጥ	<b>ሮኔሮኔ</b> ቸሮርርር	GGAGACGCCA	TCCACGCTGT	TITIGACCICC	780
	CCCCCACCCA	ጥሮሮ <u>እ</u> ਫሮሮጥሮር	CCTCGAAGCT	GATCCTGAGA	WCIICAGGGI	840
0 T 0 T 0 T T T C C	CACCCTTCAT	CUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUU	CCCTTCTTTT	CTATGGTTAA	GITCAIGICA	900
	3 C 3 3 C C T 3 8 C 3	CCCTACACAT	ATTGACCAAA	TCAGGGTAAT	TITGCALLIG	960
	* * * WCCCCCCCCC	ጥጥርጥጥጥጥልልጥ	ATACTTTTTT	GITTATCTTA	TITCIMMING	1020
	A TO COMPORT OF THE PARTY OF TH	CACCCCAATA	ATGATACAAT	GTATCATGCC	TUTTIGUALC	1020
- mmama 1 1 2 7 3	አመአ አ ሮ አ ሮሞር ል	TAATTTCTGG	GTTAAGGCAA	TAGCAATATI	ICIGCHININ	1140
		CTAACTGATG	TAAGAGGTTT	CATATIGUTA	WINGCAGCIN	1200
	***********************	ፈተጥተው ይተምተውጥ ል	TGGTTGGGAT	AAGGCTGGAT	TWITCIGNGI	1260
	COCOMPORT	ል ል ጥር ል ጥር <b>ጥ</b> ጥር	ATACCTCTTA	TCTTCCTCCC	MCMGCICCIO	1320
GGCAACGTGC	TGGTCTGTGT	GCTGGCCCAT	CACTTTGGCA	AAGAATICIA	CATCCCCCTG	1380
	, acmecemeem	- დუტტცტტნტ	GAGCTGGACA	AGIGGGAGAA	GWICEGCCIO	1440
	,	- ペケンペンスはいかは	AAGCACATC	, 101000	CCGCGAGCIG	1500
GAGCGCTTCC	CCGTGAACCC	CGGCCTGCTG	GAGACCTCCG	AGGGCTGCCG	CTACAACACC	1560
GGCCAGCTG	AGCCCTCCCT	GCAAACCGGC	TCCGAGGAGC	ACCACACCACA	GENGGCCCTG	1620
ATCGCCGTG	TGTACTGCGT	GCACCAGCGC	ATCGACGTGA	AGGACACCAA	GGAGGCCCTG	1680
GACAAGATC	G AGGAGGAGCA	GAACAAGTCC	AAGAAGAAG	T TOO A CAR A COT	CGCCGCCGAC	1740
ACCGGCAAC	A ACTCCCAGGI	GTCCCAGAAC	TACCCCATC	r cccmcaacci	GCAGGGCCAG	1800
ATGGTGCAC	C AGGCCATCTC	CCCCCGCACC	TGAACGCC.	T TOTOCCAGG	GGTGGAGGAG	1860
AAGGCCTTC	T CCCCCGAAGT	CATCCCCATC	TICTCCGCC	_ lGiccondoc	CGCCACCCCC	1920
CAGGACCTG	A ACACCATGC	r GAACACCGTG	macco caca	TAGGCCGCCAT	GCAGATGCTG	1980
AAGGAGACC	A TCAACGAGG	A GGCCGCCGAC	TGGGACCGC	C ACAMCGCCG	GCACGCCGGC	2040
CCCATCGCC	C CCGGCCAGA	r GCGCGAGCCC	COCCOCTCC	C CCATCCCCG	CACCACCTCC	2100
ACCCTGCAA	G AGCAGATCG	G CTCGATGACC	CACAACCCCC	C CCATCTACT(	r GGGCGAGATC	2160
TACAAGCGC	T GGATCATCC	r GGGCCTGAAG	AAGAICGIG	C ACTACCTGG	C CCCCACCTCC A CCGCTTCTAC	2220
ATCCTGGAC	A TCCGCCAGG	G CCCCAAGGAG	3 CCCTTCCGC	A ACTROGRATICAT	A CCGCTTCTAC	2280
AAGACCCTG	C GCGCCGAGC	A GGCCTCCCA	J GAGGIAAAG	A ACCCCCTGG	C CGAGACCCTG	2340
CTGGTGCAG	A ACGCCAACC	C CGACTGCAA	G ACCAICCIG	E CCCCCGCCC	G CCCCGGCGCC	2400
ACCCTGGAG	G AGATGATGA	C CGCCTGCCA	G GGCGIGGGC	A CCATCATGA	A CAAGGCCCGC	2460
GTGCTGGC	G AGGCCATGT	C CCAAGTCAC	C MACCCCGCC	T CCGCCAAGG	T CCAGAAGGGC	2520
AACTTCCG	A ACCAGCGCA	A GACCGIGAA	G IGCIICAAC	A AGTGCGGCA	A GGGCCACATO	2580
GCCAAGAA	T GCCGCGCCC	C CCGCAAGAA	a yactaciac	G GGAAGATCT	A GGAGGGCCAC	2640
CAGATGAA	AG ATTGTACTG	A GAGACAGGC	T WWITIIII	C CAACAGCCC	G GCCTTCCCAC	2700
AAGGGAAG	GC CAGGGAATI	T TCTTCAGAG	T CCCCCCCC	ACCAGGAGC	C ACCAGAAGAG	2760
AGCTTCAG(	ET TTGGGGAAG	A GACAACAAC	T CUCTUICAC	TA CCCACCCC	C GATAGACAAC	2820
GAACTGTA'	TC CTTTAGCTT	C CCTCAGATC	A CICITION	'Y GCGUCCC	C GTCACAATA	

## 26/29

AGATCGGTGG	CCAGCTGAAG	GAGGCCCTGC	TGGACACCGG	CGCCGACGAC	ACCGTGCTGG	2880
AGGAGATGAA	CCTGCCCGGC	CGCTGGAAGC	CCAAGATGAT	CGGCGGCATC	GGCGGCTTCA	2940
TCAAAGTCCG	CCAGTACGAC	CAGATCCTGA	TCGAGATCTG	CGGCCACAAG	GCCATCGGCA	3000
CCGTGCTGGT	GGGCCCCACC	CCCGTGAACA	TCATCGGCCG	CAACCTGCTG	ACCCAGATCG	3060
GCTGCACCCT	GAACTTCCCC	ATCTCCCCCA	TCGAGACCGT	GCCCGTGAAG	CTGAAGCCCG	3120
GCATGGACGG	CCCCAAAGTC	AAGCAGTGGC	CCCTGACCGA	GGAGAAGATC	AAGGCCCTGG	3180
TGGAGATCTG	CACCGAGATG	GAGAAGGAGG	GCAAGATCTC	CAAGATCGGC	CCCGAGAACC	3240
CCTACAACAC	CCCCGTGTTC	GCCATCAAGA	AGAAGGACTC	CACCAAGTGG	CGCAAGCTGG	3300
TGGACTTCCG	CGAGCTGAAC	AAGCGCACCC	AGGACTTCTG	GGAGGTGCAG	CTGGGCATCC	3360
CCCACCCCGC	CGGCCTGAAG	CAGAAGAAGT	CCGTGACCGT	GCTGGACGTG	GGCGACGCCT	3420
ACTTCTCCGT	GCCCCTGGAC	AAGGACTTCC	GCAAGTACAC	CGCCTTCACC	ATCCCCTCCA	3480
TCAACAACGA	GACCCCCGGC	ATCCGCTACC	AGTACAACGT	GCTGCCCCAG	GGCTGGAAGG	3540
GCTCCCCCGC	CATCTTCCAG	TGCTCCATGA	CCAAGATCCT	GGAGCCCTTC	CGCAAGCAGA	3600
ACCCCGACAT	CGTGATCTAC	CAGTACATGG	ACGACCTGTA	CGTGGGCTCC	GACCTGGAGA	3660
TCGGCCAGCA	CCGCACCAAG	ATCGAGGAGC	TGCGCCAGCA	CCTGCTGCGC	TGGGGCTTCA	3720
CCACCCCCGA	CAAGAAGCAC	CAGAAGGAGC	CCCCCTTCCT	GTGGATGGGC	TACGAGCTGC	3780
	GTGGACCGTG					3840
ACGACATCCA	GAAGCTGGTG	GGCAAGCTGA	ACTGGGCCTC	CCAGATCTAC	GCCGGCATCA	3900
AAGTCCGCCA	GCTGTGCAAG	CTGCTGCGCG	GCACCAAGGC	CCTGACCGAG	GTGGTGCCCC	3960
TGACCGAGGA	GGCCGAGCTG	GAGCTGGCCG	AGAACCGCGA	GATCCTGAAG	GAGCCCGTGC	4020
	CTACGACCCC					4080
	CTACCAGATC					4140
	GGGCGCCCAC					4200
TCGCCACCGA	GTCCATCGTG	ATCTGGGGCA	AGACTCCCAA	GTTCAAGCTG	CCCATCCAGA	4260
AGGAGACCTG	GGAGGCCTGG	TGGACCGAGT	ACTGGCAGGC	CACCTGGATC	CCCGAGTGGG	4320
	CACCCCCCC					4380
TCGGCGCCGA	GACCTTCTAC	GTGGACGGCG	CCGCCAACCG	CGAGACCAAG	CTGGGCAAGG	4440
CCGGCTACGT	GACCGACCGC	GGCCGCCAGA	AGGTGGTGCC	CCTGACCGAC	ACCACCAACC	4500
AGAAGACCGA	GCTGCAGGCC	ATCCACCTGG	CCCTGCAAGA	CTCCGGCCTG	GAGGTGAACA	4560
TCGTGACCGA	CTCCCAGTAT	GCATTGGGCA	TCATCCAGGC	CCAGCCCGAC	AAGTCCGAGT	4620
CCGAGCTGGT	GTCCCAGATC	ATCGAGCAGC	TGATCAAGAA	GGAGAAGGTG	TACCTGGCCT	4680
GGGTGCCCGC	CCACAAGGGC	ATCGGCGGCA	ACGAGCAGGT	GGACAAGCTG	GTGTCCGCCG	4740
GCATCCGCAA	GGTGCTGTTC	CTGGACGGCA	TCGACAAGGC	CCAGGAGGAG	CACGAGAAGT	4800
ACCACTCCAA	CTGGCGCGCC	ATGGCCTCCG	ACTTCAACCT	GCCCCCGTG	GTGGCCAAGG	4860
	CTCCTGCGAC					4920
	CGGCATCTGG					4980
	CGTGGCCTCC					5040
AGGAGACCGC	CTACTTCCTG	CTGAAGCTGG	CCGGCCGCTG	GCCCGTGAAG	ACCGTGCACA	5100
CCGACAACGG	CTCCAACTTC	ACCTCCACCA	CCGTGAAGGC	CGCCTGCTGG	TGGGCCGGCA	5160
	GTTCGGCATC					5220
ACAAGGAGCT	GAAGAAGATC	ATCGGCCAAG	TCCGCGACCA	GGCCGAGCAC	CTGAAGACCG	5280
CCGTGCAGAT	GGCCGTGTTC	ATCCACAACT	TCAAGCGCAA	GGGCGGCATC	GGCGGCTACT	5340
	GCGCATCGTG					5400
	CAAGATCCAG					5460
	CGCCAAGCTG					5520
	GGTGGTGCCC					5580
TGGCCGGCGA	CGACTGCGTG	GCCTCCCGCC	AGGACGAGGA	CTAACACATG	GAAAAGATTA	5640

GTARARCACC ATRAGECCECT CTAGAGGATC CARCITATE GARACTEGTS GCTGGTGTGG CCCAGATCA ATTCACCCCA CCAGTGCAGG CTGCCTATCA GARACTGGTG GCTGATTCA ATTCACTCAGA TATCACTAGA CTCCACTAGT TCTCCTAGA TCCACTAGT ACCTGTTGT TTTCATATA CATTAGAGTAT TTTACTATAA AGGAACTT AACTGGGGG ATATTAGAA GGGCCTTGAG TAGAGTATT TTTACTATAA AGGAACTT ATTTCATTTG CAATGATTAA CATAAAGAAA TTTACTAAAA AGGAACTT ATTTCATTG CAATGATTAGA CATAAAGAAA TTTACATAAA AGGAACTT ATTTGACTCAG AACAGCCTG ATGCCTATGC AACAGCCTAG TAGAGACTAT TTTACATAAA AGGAACTA ACAGCCCTG ATGCCTATGC CAATTGACT AGAGGACTA TTTACACTAA AGGGATGTA TTTGAGAGCT ATTTGACACTAC TAGAACTACT TTTACACTAC TATATGTTTT AGCTCTCA CTCAGACTAC TTTCACTAC CATTTGCTA TATATGTTTT AGCTCCTCA CTCAGACTAC TTTCACTAC CATTTGCTC CCACCCCTG ACCCCCTG ACCCCCACT CCACCACTACC CACCCCCTG CCACCCCCCT CCACCACTACCC CACCCCCCCCCC							
CCCAGATCTA ATTCACCCA CCAGTGCAGG CTGCTTATA GARAGIGNTS GTAGTACCTA GGCCCACAG TATCACTAGA CTCGCTTCT TGCTGTCCAA TTTCTATTAA  AGGTTCCTTT GTTCCCTAAG TCCAACTACT AAACTGGGG ATATTATAAA GGGCCTTGGG AGGTTCCTTT GTTCCCTAAG TCCAACTACT AAACTGGGG ATATTATAAA GGGCCTTGGG TTCTGAATAT TTTACTAAAA AGGAATGTT GGAGGGTGGT GCATTTAAAA TATTCATTCAAAA AGGAATGTT GGGAAATAT CACTATACT TAAACTCAT GAAAGAAGA GGATCAATAAA AGGATAATG CACATTGGCA ACAGCCCCTG ATGCCTATGC CTTATCCTCC GAGGGTGCAA ACAGCTAATG ACCCATTGGCA ACAGCCCCTG ATGCCTATGC CTATTCCTTC GAGGGTGCAA ACAGCTAATG AGGGCTTGA TTTGGAGGTT AAAGTTTTC CCTCAGAAAAA GGATCAAGT AGGGGCTGA TTTGGAGGAT AAAGTTTTC CCTCAGAAAAA GGATCAAGT AGGGCTTCA ATCCCACTCC ATCCCCCCCC CCTCAAGAAAA GGATCCATGT TTTACGCGCA ACTCCACTTC CTCGCCTGC CTCAAGAACAC CTCATGT TTTACGCGCA GATGGTTTC CCTCGCCTGG CTTAGTTCTT CTGTTGTCTT ATAGAGGTCT ATCTCCTTGC TTCCCCCCCCC CTGAAGTGT CCTCCACTATT TTTACGCGCA GATGGTTTC CCTCGCCTGG CTGAAGTGTT CTGTGTCTT ATAGAGGTCT ACTCCACCC CACAGTGACC CGGAATCCCT TTGAGTTATC CTGTGTCTT ATAGAGGTCT ACTCCACCC CACAGTGAC CGGAATCCCT TTGAGTTATG TCTTGTCTT ATAGAGGTCT ACTCCACCC CACAGTGAC CGGAATCCCT TTGAGTTATG TCTTGTCTT ATAGAGGTCT ACTCCACCC CACAGTGAC CGGAATCCCT TTGGCCGGAA CCCCCTATTTT TTTATTTTC TAGACGTCAG GGGCACTTT TCGGGGAAAT GGGCGGGAA CCCCCTATTTT TTTATTTTC TAGACGTCAG GGGACATGT TCGGGGGAAA CACCCGTATTT TCTTTTTTTTTC TAGACGTCAG GGGACATGT TCGGGGCAAT CAATTCCACG ACTCCACACTAAAAAAAAAAAAAAAAAAA	CONTRACTOR	ATAGGCCGCT	CTAGAGGATC	CAAGCTTATC	GATACCGTCG	ACCTCGAGGG	5700
CHARGECT GECCACARG TATCACTARG CTGSCTTICT TOTICAGA AGGSTTCCTTT GTTCCTARG TCCACATACT ANAACTGGGGG ATATTATGAA AGGSTTCCTTT GTTCCTARA AAAACATT ATTTCATTG CAATGATGTA TATAAATTAT TCTGAGATAT TTTACTAAA AAAACATT ATTTCATTG CAATGATGTA CATAAATATAT TCGAAGACTA GTCTAAACCT TGGGAAAATA CACTAATCCAT GAAAGAAGGT GAGGCTGCAA ACACCTAGG CACTATGCA ACAGCCCCTG ATGCCTATCCAT GAGGCTGCAA ACACCTAGG CACTAGGCA ACAGCCCCTG ATGCCTATCCAT TCTGACATT CTTACTAAAA AGGAACTATT TTTGAAGGT ATTGGAAGAT CCTCACACAACACA		> mmc > cccc > 2	CCACTGCAGG	CTGCCTATCA	GAAAG	GC1GG1G1G	•
AGGTTCTTT GTTCCTAAR AAAAACATT ATTTCATTG CAATGAGTA TATAAATATT 554 CATCTGGATT CTGCCTAATA AAAAACATTT ATTTCATTG CAATGAGTA TATAAATATAT 554 CATCTGGATT TTTACTAAAA AGGGAATGT GGAGGTCAGT GCATTAAAA CATAAAGAAA TCCTGAATAT TTTACTAAAA AGGGAATGT GGAGGTCAGT TAAACTCCAT GAAGAGCT GTGAGAAAA CACATTAGCA ACAGCCCCTG ATGCCTATGC CTTATTCATC GAGGGTGCAA ACAGCTAAGT AGAGGCTTGA TTTGGAGGTT TTTACATAC CATTATCATT AGAGGTCAAAAAAAAAA		~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~	ጥአጥሮኔሮሞልልር	CTCGCTTTCT	IGCIGICAN	T T T C T13T TTW-	
CATCTGGATT CTGCCTAATA AAAACATT ATTICATIG GAAGTAGA CATAAAGAAA CATCTGATAT TITACTAAA AGGAATGG GGAGTCAGT GCATTAAAA CATAAAGAAA GGAGCTGAA CATCAAACCT TGGGAAAATA CACTATATCT TAAACTCCAT GAAAGAAGG CTGAGGCTGCAA ACAGCTCAAACA GAACTATCC ACAGCTTGC ATGAGGATCA CACATTGGCA ACAGCCCTG ACGCCATGC ATGAGTATC CACATTGCTA AGGGCTGCA ACAGCCCTG ACGCCCTG ACGCCATCA CTTTACATCC CACTTGCTAC TTTACAGTAC CACATTGCTA AGGGCTGA ACAGCCCTG ACGCCATCA CTCCAGAAA GGATCAACT TTTACAGCCA AGGGCTTCA ATGAGGTTC CTTCCAGCCTG ACCCCATCACC CACATGCCC CACACAGGCC CACACAGGCC CTTAGTGTTC CTTGAGAGTAC CACCTTTCCC CTGCAGCCT TTTACGCCCA GAGGGTTCT CCTGCCTGC CCACCAGCC CACACAGGCC CTGCAAGGCC TTTAGTGTTC CTGCAGCC TTTAGAGGTC ACTTGAAGAA GGAAAACAG GGGGCATGT TTGAGAGTCA TTCTTGAGACC TTCCCCCACT CACACTGACC CGGAACCCC TTGAGGCCA GTCTAGATCA ACACGTTCCC CGCACTGCC CACACAGGCC CGGAACCCC TTGAGGCCA GTCTAGATAA AATGGTTCT TAAAAACATT CAAAATATGTA TCCGCCAAAAACAC CCCTATTTTA TTTATTTTTC TAAAAACAAT GCTTCCATG CACACTGCC CGGAACCACACACACACACACACACACACACACACA		~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~	<i>ጥሮሮ</i> እእርጥእርጥ	AAACTIGGGGG	AIAIIAIGAA	GGGCCTTG	
TICTGAATAT TITACTANAA AGGANATGG GAGAGTAGG TGAAGAGCTA GTTCAAACC TGGGAAAATA CACTATATCT TAAACTCCAT GAAAGAGGT GAGGCTGCAA ACAGCTAATG CACATTGGCA ACAGCCCCG ATGCCTATGC CCTCAGAAAA GAGCTAATG CACATTGGCA ACAGCCCCG ATGCCTATGC CTCAGAAAA GAGCTAATG CACATTGGCA ACAGCCCCG ATGCCTATGC CTCAGAAAA GAGTTCAAGT AGAGGCTTCA TTTGGAGGTT TAAAGTTTCC CTCAGAATG TATTGTTTA AGCGTCACTCA GTCTCTTCC CTCAGACTCT CTCAGCCTT ACTCCACTCA GTCTCTTCC CTGAAGTGTT CCTCCAGCTT TTTACGGCGA GATGGTTTCT CTGAAGTGTT CCTTCCATGT TTTACGGCGA GATGGTTTCT CTGAAGTGTT CTGTTCTCTT ATAGAGGTCT ACTGCAGAA CGGAAAAACAG GGGGCATGGT TTGAGTGTCC TGTTGCTCT TATAGAGGTCT ACTGCAGAA CGGAAAAACAG GGGGCATGGT TTGAGTCTC TGTTGAACCC TCTCCCCCC CTCACACAGAAACAG GGGGCATGT TAGGTTAATT TAATGGTTTCT TAGAAGAGGCC CACATGGCA GACATGACA CCCCTATTTG TTTATTTTC TAAATACATT CAGAATAACAG GGGCAATGTT AGGCTAAAAA CCCCTATATAT ATATTTTTC TAAATACATT CAAATAATGTA TCGGCGAAAA CATTTCCGTG AGCAATAAC CCCCATTATT CCCCTTTTTT GCGGCATTT TCGGGGAAAT AGACAATAAC CCTGATAAAA GCTTCAAAAA TATTGAAAAA GGAAGAGATA CATTTCGTG TCGCCCTTAT CCCCTTTTTT GCGGCATTTT GCGTCACCTG ATCAACAGG ACCCCTTATAA AGATCTGGTT GCGGTCACCC AGAGGGTTAC CCAATGACGA ACTCCAACAG CGGTAACAC CTTAAGAGGTT TTTGCCCCC AGACACTTTTC CCAATGACGA AACCCATTAAA AGATCTGCTAGAGGTT TTTGCCCCCCAAACAC TATTCACCAG CGGAAACAC TATTCACCAG AGAACGTTTT AGGGCAACAC AACTCGGTCG CCGCAAACAC TATTCTCACAA AGAACCATTATAC CCAATGACGA AAAACCATC TACCGGCAACAC TATTCTCACAA AGAACCATTATAC CCAACAACAC TTTTTTCCAC CACAATGGGG ATCACCTTAACAC TATTCTCACAACAC AACAACCAC TATCTCTCAC AACACATTACC CCAGCACAACA TTTTTCCACAACAC TATACTCTCAC AACACATTACC CACCATACAC CAACAACCT TCCGCCAACAT ATAACCACC CACCATCAC CACCATCAC CACCATCAC CCAGCACACA TGCACCACAC ATTAACCAC CACCATCAC CACCACACAC CCAGCACCAC TTTTTCCACACAC AACAAACCAC CACCACTAC CACCACACC CCAGCACCAC TTTTTTCCAC CACCACACC TTTTTCACCC TATTCTCAC CACCACACAC CCAGCACCAC TTTTTTCCAC CACCACACC TTTTTCACCC TATTCTCAC CACCACACC CCAGCACCAC TTTTTTCCAC CACCACCACC CACCACCACC CCCCCACACAC CCAGCACCAC TTTTTTCCAC CACCACCC CACCACCACC TTTTACACCC CACCACCCC CCAGCACCAC TTTTTTTCCAC CACCACCCC CACCACCACC CCACCACCC CCCCACCA		AMACAMA AMA	יוייוייתראמרגגגג	ATTICATIG	CWHIGHTOTA	7 7 7 4 W T T T T T T T T T T T T T T T T T T	
TRARAGAGETA ACAGCTARAGE CACATEGGA ACAGCCCETG ARGCTRATE CITATICATE GAGGETGCAA ACAGCTARAGE ACAGCCCTG ACACCCCCTGA ACAGCTATAGE CCTCAGARAA GGATTCAAGT AGAGGCTTGA TTTGGAGGTT TTTCACTACC CATTTGCTTA TTTACATTAC TTATTGTTTT AGCTGTCCT ATGCAGTCT TTTCACTACC CATTTGCTTA TCCTGCACTC CTCAGCCTTG ACTCCACTCA GTCTCTTGC TTTTAGAGATAC CACCTTTCC CCTGAAGTGTT CTTTACGCCAT ATTACGGCGA GATGGTTTCT CCTCGCCTGG CCACCTACC TTGATGTGTCT CTGTGAGCCT TTTTACGGCGA GATGGTTTCT CCTCGCCTGG CCACCTACC TTGATGTGTCT CTGTGAGCCT TCTTCCTGC CTCCCCACT CACAGTGACC CGGAATCCCT TTGACTGTCC TGTGAGCCT TCTTCCTGAGAGT TCCGGGAAAT GTGCGCGGAA CCCCTATTTG TTTATTTTTT TAAATACATT CAAATACTAT TCCGGTCAAG GGCCAATAAA CCCCTATTTG TTTATTTTTT TAAATACATT CAAATACTAT TCCGGTCAAG GACAAATAAC CCCGATAAAA TAAAGAGTCT TGCGCTAGAAT TTTGCCCCCAA ACCCCTTATA ACCCTTTTTT GCGGCAATTT TCCGCTCTGT TTTTCCCCCAC ATCCAACAG AACTCGATCCC CGGTAACAC TATTCTCAGA GCCAATAACA GCCCTTTTAA AGTCTCTAA TGTGGCCCGG TATTATCCCG TATTGACGCC CCAGTCACAG AACTCGGTCG CCCCATACAC TATTCTCAGA ATGACTTGGT TGGTTACCAC GCCAGTCACAG AACTCGGTCG CCCCATACAC TATTCTCAGA GACCAAACGT TCCGCAAACAC TATTACCTGA AGAACATTT TCCGGAACAC GCACAAACGT TCCGCAAACAC TATTACTCTGA AGAACATTT TCCGGAACAC TTAAATAGACT GGATGAACAC TGCGGCCAC AGACACACAC TCTCGCCTGA AGAACACT TTTTCCCA CAACAATCGG AGACCACAC TCTCAGCACAC TTAAATAGACT GGATGAACAC TATAACTTGC ACACAACAC TCCCGCATACAC TTTTACACG CGCCGCAACAC ACCAAACACT TTTTCCCGC CAACAACACT TTTTCCCCG TTTTTCCCC CAACAACACT TTTTCCCGC TTTTTCCCCG CAACAACACT TTTTCCCGC TTTTTCCCC CAACAACACT TTTTTCTCC CAACAACACT TTTTCCCGC TTTTTTCCCC CAACAACACT TTTTTCATCC CAACAACACC TTTTTTCATCC CAACAACACT TTTTTCATCC CAACAACACC TTTTTTCATCC CAACAACACC TTTTTTTT		መመጠን መመን እ እ እ	ACCCA ATGTG	GGAGGTUAGT	CCWITINGS.	CULTABLOSS	
GAGGCTGCAA ACAGCTAATG CACATTGGCA ACAGCCCTG CCTCAGAAAA GGATTCAAGT AGAGGCTTGA TTTGGAGGT TATGATGTT TATTCACTAC TCTTACATTAC TTATTGTTTT AGCTGCTC ATGATGTCT TTTCACTACC CATTTGCTTA TCTGCATCT CTCAGCCTTG ACTCCACCTC GTTCTCTTCC TTTAGAGATAC CACCTTTCCC CTGAAGTGTT CCTTCCATGT TTTACGGCA GATGGTTTC CCTGCCTGG CCACTCAGCC TTAGTTGTCT CTGTGTCTT ATAGAGGTCT ACTGAAGAA GAAAAACAG GGGCATGGT TTGACTGCC TGTGAGCCT TCTTCCCTGC CTCCCCCACT CACAGTGACC CGGAATCCCC GGCATGGCA GTCTAGATAT ATGAGAGTCT ACTGAAGAG GGAAAAACAG GGGCATGGT TTGACTGCC TGTGAGACA TTCTTCAAGA CGAAAGGCC TCGTGATACG CCTATTTTTA TCAGCTGCAC GTCTAGATAT ATGTTTATTTC TAAATACATT CAAATAATAT TCCGCTCATG GTGCGCGGAAAC CCCCTATTTG TTTATTTTC TAAATACATT CAAATAATAT TCCGCTCATG AGACAATAAC CCCCTATTTG TTTATTTTC TAAATACATT CAAATAATAT TCCGCTCATG AGACAATAAC CCCGTAAAAT AAAAGATCCT TGCGCCGAA ACTCCACAACAC CATTTCCTTC TCGCCCATT TCCCTTTTTTT TCCCACCCGA AGACGTTTC AACTCCACACAC CAATCAATAA AAAAGATCCT TTGCGCCCGA AGACGTTTC AACTCCACACAC CAATCAATCA AACTCCGGTCACAC CAACACAC CAACACACAC CAACACACAC CAACACACAC CAACACACAC AACTCCGGCCACAC TAACACTCAACAC CAACACACAC CAACACACAC		<b>ማመጠጣ እ እ አ ጦርጥ</b>	TCCCAAAATA	CACTATATUT	IMMULTICAL	Old High and a	6060 -
CCTGAGAAAA GGATTCAAGT AGAGGCTTGA TITGGAGGT AGAGTTCCC TTTACATTAC TTATTGTTTT AGCTGCCCCAC AGAGGTCTC TTACACTAC CACTTGCCC TCCGCATCT CTCACCCTT ACTACGCTA GTTCTCTTGC TTAGAGATAC CACCTTGCCC CTGAAGTGTT CTTACCGCCAC ACTCCACCTA GTTCTCTTGC TTAGAGATAC CACCTTGCCC CTGAAGTGTT CTTTCCATGT TTTACGGCGA GATGGTTTCT CCTCGCCTGG CCACTCAGCC TTAGTTGTCT CTGTGAGCCCT TCTTCCCTGC CTCCCCCACT CCCCCCTG CCACTCAGCC GACATGGCA GTCTAGATCA TTCTTGAAGA CGAAAGAGA GGAAAAACAG GGGGCATGGT TAGAGTTAATG TCATGATAA ATGGTTTCT TAGACGTCAG GTGCCACTT TCGGGGAAAT AGGGTAAATG TCATGATAAA TATGGTTTCT TAGACGTCAG GTGCCACTT TCGGGGAAAT AGGCAATAAC CCTGATAAAT TATTGTTTTTC TAAATACATT CAAATATGTA TCCGCTCAG GTGCCACTTT TCGGGGAAAT AGACAATAAC CCTGATAAAT GCTTCAATAA TATTGAAAAA GGAAGAGTAT TCCGCTCATG AGACAATAAC CCTGATAAAT CCCTTTTTTT GGGGCATTT TCCCCTCAT TTTTTCCTCAC CCAGAACCC TGGTGAAAGT AAAAGAATCC TAAATAGAAT TCCCCTTTTTT TCCCCCCCAC AGTGGGTTAC CCAGAACCC TGGTGAAAGT AAAAGAATCC TTGAGAGTT TTCCCCCCACA AGAACGTTTCAA AGCAATGATGA GCACTTTTAA AGTTCTGCTA TGTGGCGCGG TATTATCCCC CCAGTCACAC AAAAGCATCT TACGGATGGC TATTATCCCC TATTGACCAC ATAACCATGA GCACTTTTTAA AGTTCTGCTA TTGTGGCCCGA AGAACGTTTT AGAGCACACAC TACCACACACAC TATTCTCAGA ATGACTTGGT TGGTGCACG AAAACCAT TCCGGCCACAC TATCTCTCAGA ATGACTTGGT TGGTTCACAC CCAGACACCA AGAACGATC TACGGATGGC GACCACTACAC TATCTCTCAGA ATGACTTGGT TGGTGCACG AAAACCAT TCCGCCCAACAC TATCTCTCAGA ATGACTTGCT TCGTTGGAAA CCCGGACACAC TGCGCCACAC TATCTCTCAGA ATCACCTTCA ACCAAACCAC CCGACACACA TGCGCACACAC TATCTCTCAGA ATCACCACACACACACACACACACACACACACACACAC		- C - C C C C - 3 C C C	ぐれぐれががににじる	ACAGCCCCCCC	VICCIVIO	CITITITITI	
TTACATTAC TTATGETTT AGCTGTCCTC ATTACTCTC TTAGAGATAC CACCTTCCC 630 CTGCAGCATC CTCAGCCTTG ACTCCACCTC AGTCCTCTGC TTAGAGAGATC CACCTTCCC 651 CTGAGAGGTT CCTGCAGCTT TTAGAGGATA CACCTTCCC 651 CTGAGAGTGT CCTGCAGCTT TTAGAGGATA CACCTTCCC CTCCCCACT CACAGTGACC CGAATCCCT TTAGACTGTC TGTAGACCCT TCTTCCCTGC CTCCCCACT CACAGTGACC CGAATCCCT TTAGACTGTC TGTAGACCCT TCTTCCTGC CTCCCCACT CACAGTGACC CGAATCCCT TTAGACTGTC TGTAGAGACTA TTCTTGAAGA CGAAAGGGC TCGTGATAACG CGAAACGACT TCCGGCAGAACGC TCATATTTA TAGACGTCAG GTGGCACTT TCCGGCAGAACGC CCCTATTTT TAGACGACAC GTGGCACTT TCCGGCAGAACGC CCCTATTTT TAGACGACAC GTGGCACTT TCCGGCAGAACGC CCCCTATTT TTATTTTT TAGACGACAC GAACAGATAT CAAATATGATA CACTTCCAGACACAC ACTCCAACAACACAC CACCAGAACACC TAGACACACACACACACACACACACACACACACACACACA			አሮአርርርጥጥርል	TYPYTY CARCICLE	AAAGIIIIGC	TWIGGIGIA	-
TCCTGCATCT CTCAGCCTTG ACTCCACTCA GTTCTCTTC TTAGGARTAC CTGAAGTGTT CCTTCCATGT TTTACGGCGA GATGGTTTCT CCTGGCCTGG			አሮሮጥርጥሮሮሞሮ	ATGAATGICI	TITCACIACC	CHILLOCALL	_
TTGATGTTC CCTTCCATGT TTTACGGCCA GATGGTTTC CCTGCATGC GGGCATGGT TTGATGTGTCT TATAGAGGCCT TCTTCCCTGC CTCCCCACT CACAGTGACC CGGAATCCCT GACAGTGACC CGGAATCCCT CACAGTGACC CTGCACAGGCC TCTTCACAGG CCTACTTTTA CACAGTGACC CTGCACAGGCC TCGCACAGGCC CGGAATCCCT CACAGTGACC CGGAATCCCT CACAGTGACC CTGCACAGGCC CTGCACAGGCCC CCCACAGAACAC CCCCATATAAT TATTGAAAAA GGAAGAGTAT CACAGAACAGC CTGCACAGAACAC CTTCACAGAC CTCCACAGAACCC CCCACACACAC ACCCACACAC CTTCACAGGC CCCACACACAGAACACAC ACCCACACAC TATTCCCACAG CCCACACACAC ACCCACACAC CTTCACACAG CCCACACACAC ACCCACACAC CTTCACACAG CACACACACAC ACCACACACAC ACCACACACA		~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~	አርጥርር እርጥር A	GTTCTCTTGC	LINGUGUTUC	Cuccilian	-
TTAGTTGTCT CTGTTGTCTT ATAGAGGTCT ACTIGARGAS CANAMARAN CGARATCCT TTGACTGTCC TGTTGAGCCCT CTCCCCCACT CACAGTGACC CGGAATCCCT GCACATGGCC TGTTAGACCA TTCTTGAACA CGAAAGGCC TCGTGATACG CTATTTTA 65.  GTGCGCGGAA CCCCTATTTG TTTATTTTC TAAATACATT CAAATATGTA TCCGGTCATG GGCACATAAA ATGGTTTCT TAAATACATT CAAATATGTA TCCGGTCATG GAGACAATAAC CCTGATAAAT GCTTCAATAA TATTGAAAAA GGAAGATAT GAGTTACCAC AGACAATACAC CCTGATAAAT CCCCTTTTTT GCGCGCATTT TCGGCCCGA AGACGTTAC CAATTCCAG ACCCCTATT TCCCTTTTTT GCGCGCATTT TCGGCCCGA AGACGTTAC AGACAATCAC TCCGATAAACAC CGGTAAGATC CTTGAGAGTT TCCGCCCGA AGAACGTTT AGATCACAC TATTCCACAC AAATACCAC AGACCATTTAA AGACCACAC AACTCCCAC TATTCCACAC AACTCCCCCAC TATTCCACAC AGACCATTTT AGACCACAC AACTCCACAC AACTCCACACACACAC ATGACCACAC AACTCCACACACACACACACACACACACACAC		COMMOC N TOT	TOTO TOTO TOTO TOTO TOTO TOTO TOTO TOT	GATGGTTTCT	CCICGCCIGG	CCNCTCMCCC	
TTGACTGTCC TGTGAGCCCT TCTTCCCTGC CTCCCCCACT CACGAGGACC CGAATGGCA GTCTAGATCA TTCTTGAAGA CGAAAGGGCC TGGTGATAGG CCTATTTTA 65 TAGGTTAATG TCATGATAAT AATGGTTTCT TAGACGTCAG GTGGCACTTT TCGGGGAAAT GTGCGCGGGAA CCCCTATTTG TTTATTTTC TAAATACATT CAAAATAGTA TCCGCTCATG AGACAATAAC CCTGATAAAT GCTTCAATAA TATTGAAAAAA GAAAAAACC ACCGCTACCA CATTTCGGTG TCGCCCTTAT TCCCTTTTTT GCGGCATTTT GCGCTCATG TTTTGGTCAC CCAGAAACGC TGGTGAAAGT AAAAGATGCT GAAGATCAT TGGGGTGCACG AGTGGGTTAC CCAATGATGA GCACTTTAA AGTCTCGCTA TGGGGGCACT TTTGGCCCCGA AGAACGTTT GCGCAATGATGA GCACTTTAA AGTCTCGCAC TATTCCCGT TTTTGCTCAC GGGCAAGAGC AACTCGGTCG CCGCATACAC TGTGGCGCCGC TATTATCCCG TATTGACCCC CCAATGATGA GCACTTTTAA AGTCTCGCAC TTTCTCAGA ATGACTTGT TAACTAGACGC CCAGTCACAG AAAAACACT TACGGATGGC ATGACAGTAA GAGAATTATG CACTGCTCCA ATAACCATCG GTGATAACAC TGCGGCCAAC TTACTTCTGA AGAACTTGT TAACTGCGCCCAC ATAACCATCG GTGATAACAC TGCGGCCAAC TTACTTCTGA CAACGATCGC AGACACTACAC CAACAACGC GTGATAACAC TGCGGCCAAC TTACTTCTGA CAACGATCGC AGGACCGAAG CCAGTCAACG ATGAAGCCAT ACCAAACGAC GAACCATTAC CTCAGCTTCC TGCTGCGCAAC CCAGACAACAC TGCGCAAACT ATAACTGGC GAACCACTAC TCCTAGCTTC CCGGCAACAAC TTAATAGACCT GGGTGAAACT TAAACTGGC GAACCACTAC TTCTCGCTC GGCCCTTCCG TTACTTCAC GCGCAACAACT TTAATTGCTCA AAAATCACAC GCAGAACAC TTCTCGCTC GGCCCTTCCG TTAATTGCTCA AAAATCACAC GCAGGAACAA TGCAGAACAC AAAATGACAC ACCGAGACCA TTCTCGCTC CGGTAACACT TTAATTCAAC GCCGGAACAA TGCAGAACAC AAAATGACAC ACCGCTACAC TTCTCGCTC CGGTAACCT TTTTTAATTTA AAAGCATCA AAAATGACAC ACCGCTGAAC AAAAAAACC ACCGCTACAC TTCTCGCTC GCGTAATCTT TTTTTCTGCG CGTAATCTC TTTTTTGATA TAACTGC CAACACTTT TTTTTCTGCG CGTAATCTC TTTTTTTTTT			አጥአሮአሮርጥርጥ	ACTITICAAGAA		99999711	
CGACATGGCA TAGGTTAATG TCATGATAAT AATGGTTTCT TAGACGCCGGAA CCCCTATTTG TCATGATAAT AATGGTTTCT TAGACGCCGCACTT TCCCCCTATTTG TCATCACAAC CCCCTATTTG TCATCACACAC CCCCTATTTG TCACCCCTATTT TCCCCTTTTTT TTAATACACTT CCAGACACCAC CCAGAAACGC CCAGACACCAC CCAGACACCAC CCAGACACCAC CCAGACACCAC CCAGACACCAC CCAGCCCCTAT CCCCTTATT AAAGACCACAC CCAGCCCCTTAT TCCCCCTTATT TCCCCCTTTTTT CGCCCCTTAT TCCCCCTTTTTT CGCCCCTTAT TCCCCCTTTTTT CGCCCCTTAT TCCCCCCTTAT TCCCCCTTTTTT CGCCCCTTT TTTCCCCCCCA AGAACGCT CCAGACACCG CCAGTCACCAC CCAGTCACCAC CCAGTCACCAC AACCCGCCCCC CCAGTCACCAC AACACCGCCC CCAGTCACCAC AACACCGCCC CCAGCCCCCC CCAGCCCCCC CCACCCCCCCC			<i>-</i> መሮጥሞሮርር	CTCCCCCACT	CACAGIGACE	CGGUTTICCCT	
TRAGGTTANTG GTGCGCGGAA CCCCTATTTG TTTATTTTC TAAATACATT CAAATATGTA GGAGGAATAAC CCTGATAAAT GCTTCAATAA TATTGAAAAA GGAAGGATAT GCATTCCCTG TCCCCTATT TCCCTTTTTT GCGCCCTTT TCCCCTTTTTT GCGCCCTTT TCCCTTTTTT GCGCCCTTT TCCCCTTTTTT GCGCCCTTT TTTTGCTCAC AGACAACGC CTGGTGAAAGT AAAAGATGCT TGGGTGCACG AGTGGTTAC AGACACTGGTCG ACCATGACACA CCCAATGATGA GCACTTTTAA AGTTCTCACAC CCCAATGATGA GCACTTTTAA AGTTCTCACAC CCCAATGATGA CCCAGTCACA AAAACCATC TACGGATGCC CCAGTCACA ATGAAGCCAT ACCAAACGAC CGACCACTGA CCCGGACCTGA ATGAAGCCAT ACCAAACGAC GCACCACTGG CCGGACCACA TTATTCTCAG AACACCG CCGGACCACA TTATTCTCAG CCAGCACCT TTATTGCTCA AACAACGAC GCAGCACCT TTATTGCTCA AAAACCAC CCAGACCAC TTCTCGCCTTGA TCGTGGAAC CCCGGACCAA CCACAACGAC CCCGGACCAA CCACAACGAC CCCGGACCAA CCACACCC CCGGCCAACA CCCGGACCAA CCCCGACCAC CCCGCACCAC CCCGCCCCC CCGCCCCCC CCGCCCCCC CCGCCCCCC			T	( (JAAAGGGCC	ICCIDATACA	CC 1441 4 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	
GTGCGCGGAA CCCCTTATTG TTTATTTTT TAAATACATT CAAATACATT GAGAGAGAAAA GGAGAGATAA GAAAAAAAAAA	<del>-</del>		ա <b>Հաևաստաստաստա</b>	Transaction LAG	(コー(コ(コ) ハ) コートー	1000000	
AGACAATAAC CCTGATAAAT TATTGAAAAT GGACGACTTC CATTCCGTC TCGCCCTTAT TCCCTTTTTTT GCGGCATTTT GCGTCCCTTAT TTTTGCTCAC 67 CCAGAAACGC TGGTGAAAGT AAAAGATCCT GGAGACACT TGGGTGCACG AGACGTTTT 68 CCAATGATGA ACTCCAACAG CGGTAAGATC CTTGAGAGTT TCGCCCCGA AGAACGTTTT 69 CCAATGATGA GCACTTTTAA AGCTCCGTA TGTGGGGGGG TATTATCCCG TATTGACGCC 69 CCAGTCACAG AAAAGCATCT TACGGATGGC ATGTCACAG AAAAGCATCT TACGGATGGC ATGTCACAG AAAAGCATCT TACGGATGGC ATGTCACAG AAAAGCATCT TACGGATGGG GATCATGTAA CACGAACAGA ATGACTTGGT TGGGTGCACG AGGACCGAAG 71 CCGGACCATGA GTGATAACAC TACGGATGGG GATCATGTAA CTCGCCTTGA TCGTTGGGAA 72 CCGGACCTGA ATGAAGCCAT ACCAAACGAC GACCTACTTA CTCTGCCTC TGGGTACACA ACCAAACGAT TACCACTGAA CACAAACGAC CACGATGCC CCACGACAA 72 CCGGACCACGT TTATTGCTGA ACCAAACGAC GACCGTGACA TCCTGGCTC GGCCCTTCCG TTATATTGCTGA TAAATCTGGA GCAGGACCAC TTCTGGCTC GGCCCTTCCG TTATATTGCTGA TAAATCTGGA GCAGGACCAC TTCTGGCTC GGCCCTTCCG TAAAACCAC GACCACAGATGC TTATATTGCTGA TAAATCTGGA GCAGGACACT TTCTGGCTC GGCCCTTCCG TTATATTGCTGA TAAATCTGAA CCCGGAACAA TTATATCTTA TAAATCTGAA ACGCCTCC CCTATCGTAG TTATCTACAC GACGGGGAGT TTTTTATTTTA			A TATAL TO THE PERSON AND A STATE OF THE PERSON AND A STATE OF THE PERSON AND ADDRESS AND	TAAATACATT	CAAAIAIGIA	10000	6720
CATTTCCGTG TCGCCCTTAT TCCCTTTTTT GCGGCATTTT GCGTGCACG AGTGGGTTAC 68 CCAGAAACGC TGGTGAAAGT AAAAGATGCT CTTGAGAGTT TTGGCCCCGA AGAACGTTTT 69 CCAGTAGATG ATCTCAACAG CGGTAAGATC CTTGAGAGTT TTGGCCCCGA AGAACGTTTT 69 CCAATGATGA GCACTTTAA AGTTCTGCTA TGTGGCGCGG TATTATCCCG TATTGACGCC 69 CCAGTCACAG AAAACCATCT TACGGATGC ATGACACTAA GAGAATTATG CACGATCGC CAGTCACCA AAAACCATCT TACGGATGCC ATGACATAACCATGA AGAACGATCGA ACAACGATCGG AGGACCGAAG AAAACCATCA ACAACGATCGG GACCGAAG ACCAACGATCGG CAACAACGA CCACGATGCC CAACAACGA CCACGATGCC CAACAACGA CCACGATGCC CAACAACGA CCACGATGCC CAACAACGA CCAACAACGA CCACGATGCC CAACAACGA CCAACAACGA CCAACAACGA CCAACAACGA CCAACAACGA CCACGATGCC CCGGCCAACA ACCAACGATCG GGCCCAACA ACCAACGATCG GGCCCAACA ATGAACCAA ACCAACGATCG GGCCCATCCG CAACAACGA CCACGATGCC CCGGCAACAA ACCAACAGCC GGCCCATCCG GGCCCTTCCG CAACAACGA CCAACAACGA CCACGATGCC CCGGCAACAA TATTACTGCA GGACGAACA TTATCTCACAC GGCCAACAC TATTACTCACA GGCCGCAACA ACCAACACGT CCACGATGCC CCGGCAACAA TATTACTTGCA GACCAACACT TAATCTCACAC AAATACACAC ACCGGTACCA TATTACTCACAC AAATACACAC ACCGCTACCA TATTACTACAC ACCGGGAAGT TAACGTGAACA ATCGCTGAGA TAACGTGAGT TATTCTACACC ACTGATTAACT AAAACACTCA TATTATACTTT AGATTGATTA AAAACATCCT TTTTCTTCGC CGTAATCTCC TATTTTGATA ACCGCTACAC ACCGCTACCA TACACCAACAC TACCGCTACAC ACCGCTACCA TACACCAACAC TACCGCTACAC ACCACTTC TTTTCCGCG TACATCTC TACACCAAC ACCGCTACCA TACACCACAC CCAACACCTT TTCCGAACCA ACCGCTACCA CAAAAAAAACC ACCGCTACCA TACACACAC CCAACACCTT TTCCGAACCA ACCGCTACCA TACACACACAC CCAACACCTT TTCCGAACCA ACCGCTACCA CAAAAAAAACC ACCGCTACCA TACACACACAC TACACACAC CCAACCTTC TACACCGAAC TACACACAC CCAACCTTC TACACCACAC TACACACAC			ם ביויב ביויחידים	TATTICABABA	COMMUNICATION	O170 1111 1 01-1	6780
CCAGAAACGC TGGTGAAAGT AAAAGATGCT GAAGATC.AGT TGGGCCCGA AGAACGTTTT 69 ATCGAACGG ATCTCAACAG GCGTAAGATC CTTGAGAGTT TTCGCCCCCGA AGAACGTTTT 69 CCAATGATGA GCACTTTTAA AGTTCTGCTA TGTGGCGCGG TATTTATCCCG TATTGACGCC 69 GGGCAAGAGC AAACGCATCT TACGGATGC ATGACAGTAA AGAACGTTGT TACGGAAACGA TACTCTCAGA ATGACTTGGT TGAGTACTCA 70 ATAACCATGA GTGATAACAC TGCGGCCAAC TTACTTCTAGA CAACGATCGG AGACGAAGG CCGGAGCTGA ATGAACCAT ACCAAACGAC GAACGATCGG GAACCGAACG CAACAACGCT TGCGCCAACCA ATGACATCATA CTCGCCTCGA TGTAGGAAACGAC CAACAACGAC CAACAACGATCG GAACCAACGAT TCTTAATAGACT GGAACGAACGA ATTAACTGGC GAACTACTTA CTCTCAGGCTC CGGCCAACA ATTAACTGGC GAACTACTTA CTCTCAGCTC CGGCCAACA 71 GCAGCAACGT TGCGCAAACT ATTAACTGGA GAACGATCACT TCTTGCGCTC GGCCCTTCCG 71 GCAGCAACTA TGGATGAGC GAATACACA ATTAACTGGA GACGGAACAC TTCTTGCGCTC GGCCCTTCCG 71 GCAGCAACTA TGGATGAACG AAATACACAC ACCGGTGACA ATTAACTTGCA GACGGACCAC TTCTTGCGCTC GGCCCTTCCG 71 GCAGCAACTA TGGATGAACG AAATACACAC ACCGGTGACC TTCTTGCGCTC GGCCCTTCCG 71 CATTGGTAAC TGTCAGACCA AGTTTACTCA TATAACTTTA AGAGTCTCA ACTGATTAAC TTATAACTTTA AAACGTTCAT TTTTCTTCTGCG GGTGAAGATC CTTTTTGATA ACTCTCATGAC CAAAATCCCT 71 TAACGTGAGT TTTCGTCCA CTGAGCGTCA GACCCCGTAG AAAAAAAACC ACCGCTACCA 71 AGCAGAGCCC TTCTTCTCCC CTGAGCGTCA CAAACACCA ACCGCTTCC 71 AGCAGACCCT TTTTTCTCGCG CGTAATCTCC TGCTTGCAAA ACCCCCTTC TTCCGAAGGT ACCCCCTTC TTCTCCGAAGCT ACCCCCTTC TTCTCCGAAGCT ACCCCCTTC TTCCCGAAGCT ACCCCCTTC TTCCCGAAGCT ACCCCCTTC TACACCAC CCAACTCTT TTCCCGAAGCT ACCCCCTTC TACACCAC CCAACTCTT TTCCCGAAGCT ACCCCCTTC TACACCAC CCAACTCTT TTCCCGAAGCT ACCGCTTCC TACACCACC TACATCCCT TTCCCGAAGCT ACCGCTTCC ACACCCTTC TTCCCGAACG ACACGCTTCC TACACCACAC TTCCCAACCCC TACACCCTC TTCCCAACCCC TACACCCTC TTCCCAACCCC TACACCCTC TTCCCAACCCC TACACCCTC TTCCCAACCCC TACACCCC TACA			, ա,	GCGGCATTT	GCCTTCCTGT	1111001010	6840
ATCGAACTGG CCAATGATGA CCAATGATGA CCAATGATGA CCAATGATGA CCAATGATGA CCCAATGATGA CGCCATTATAA AGTTCTGCTA TGTGGCGCGG TATTATCCCG TATTGACGCC GGGCAAGAGA CAACTCGGTCG CCGCATACAC TATTCTCAGA AAAGCATCT TACGGATGGC AAAAGCATCT ATCGGATGGC ATGACAGTAA CAACATCGG GGGCAACAC ATAACCATGA GTGATAACAC TGCGGCCAAC CTTTTTTGCA CAACATGGGG GACCAACACGT CCGGAGCTGA ATGAAGCCAT ATCACACGAC CCGGAGCTGA ATGAAGCCAT ATCACACGC CCGGAGCTGA ATGAAGCCAT TGCGCAAACT TGCGCAAACT TTAATAGACT GGATGGAGC GGATAAAGTT CCGGCAACAC TTAATTGCTGA GCAACAACGT GGATGGAGC CAACAACGGC GGATAAAGTT GCAGGACCAC TTAATTGCTGA GCAGCACTG GCAGCACTG GCAGATGG CAGCACTG GCAGCACTG GCAGCACTG GCAGCACTG GCAGCACTG GCAGCACTG CAGCACTG GCAGCACTG GCAGCACTG GCAGCACTG GCAGCACTG CAGCACTG GCAGCACTG CAGCACTG CAGCACTG CAGCACTG CAGCACTG CAGCACTG CAGCACTG CAGCACTG CAGCACTC CATTGGTAC CATTGGTAC CAATGGTAC CATTGGTAC CAATGGTAC CAAAACCCT CAAAACCCC CAAAACCCC CAAAACCCC CAAAACCCC CAAAACCCC CAAAACCCC CAAAACCCC CCGAACCA CAAAAACC CAAAAACC CAAAAACC CAAAAACC CAAAAACC CAAAAACC CAAAAACC CAAAACCCC CAAAACCCC CAAAACCCC CAAAACCCC CCGCACCAC CTTTTTGATA CCACACAC CCACACTTC CCGCAAGGG CCCCCCCC CCGCACCAC CCACACTC CCGCCACCCC CCGCACCAC CCACACTC CCGCCACCCC CCGCCACCCC CCGCCACCCC CCCACCACTC CCCCCCCCC CCGCCACCCC CCCCCCCC CCGCCCCCC CCCCCCCC			• *************************************	CAACATCACT	114001100い	MO100011110	6900
GGCAAGAGC AACTCGGTC CCGCATACAC TATTCTCAGA ATGACTTGGT TGAGTACTCA 70 CCAGTCACAG AAAAGCATCT TACGGATGC ATGACAGTATAG CAGGGCAAGAG ATAACCATGA GTGATAACAC TGCGGCCAAC TTACTCTCAGA ATGACTTGGT AGGACCGAAG 71 ATAACCATGA GTGATAACAC TGCGGCCAAC TTACTTCTGA CAACGATCGG AGGACCGAAG 71 CCGGAGCTAACCG CTTTTTTGCA CAACATGGG GATCATGTAA CTCGCCTTGA TCGTTGGGAA 72 CCGGAGCTGA ATGAAGCCAT ACCAAACGAC GAGCGTGACA CCACGATGCC TGTAGCAATG 72 CCGGAGCTGA ATGAAGCCAT ACCAAACGAC GAGCGTGACA CCGCCAACACA 73 CCAACAACGT TGCGCAAACT ATTACTGGC GAACTACTTA CTCTAGCTTC CCGGCAACAA 73 CCAACAACGT TGCGCAAACT TATACTGGA GCCCGTCAGC TTCTGGCTC GGCCCTTCCG 73 CCAGCACTGG GGCCAGATGG TAAACCTCGA GCCGGTGAGC TTCTGGCTC GGCCCTTCCG 73 CCAGCAACTA TGGATGAACG AAATAGACCAG CCGTGAGA TAGGTGCCTC CGGCAACAA 74 CCAGCAACTA TGGATGAACG AAATAGACCAG ATCGCTGAGA TAGGTGCCTC CGGCAATCAG 74 CCAGCAACTA TGGATGAACG AAATAGACCAG ATCGCTGAGA TAGGTGCCTC CAAAATCCCT 74 CCAGCAACTA TGGATGAACG AAATAGACAG ATCGCTGAGA TAGGTGCCTC CAAAATCCCT 74 CCAGCAACTA TGGATGAACG AAATAGACAG ATCGCTGAGA TAGGTGCCTC CAAAATCCCT 74 CAACATCCTT TTTCGTTCCA CTGAGCGTCA GACCCCGTAG AAAAGAACA ACCGCTACCA 75 CCGGTGGTTT TTTCTTCGC CGTAATCTGC TGCTTGAAA AAAACATCAA ACCGCTACCA 75 CCGGTGGTTT TTTCTTCGC CGTAATCTGC TGCTTGAAA AAAACATCAA ACCGCTACCA 75 CCGGTGGTTT TAGCACCAC TCAACACTCTT TTCCGAAGGT AACTGGCTTC 75 CCGCAGGGCG AAAACCCCA TACATACCTC CTAGTGTAGC CCACACTTC 75 CCGCAGGGCG CGGCCTGAAC GGGGGGTTCG TCCGGAAGAAA ACC CAACATCCT 75 CCACGCTGCC TTCACCGG TTGGACCACAC CCACACTTC 75 CCACGCTGCC TTCACCGG TTGGACCACAC CCACACTTC 75 CCACGCACGC CCGGCCTGAAC GGGGGGTTCG TCCCGAAGGG CCACACGC CCACGCT TCCCGAAGGG CCACACGC CCACACCCC CCACACGC CCACACGC			• ^^^^	CTTGAGAGTT	TILGCCCCC	MOINICOT	6960
GGGCAAGAGC CCAGTCACAG AAAAGCATCT TACGGATGCC ATACACTAGAG AAAAGCATCT TACGGATGCC ATACACTAGAG AAAAGCATCT TACGGATGCC ATACACTAGAG AAAAGCATCT TACGGATGCC ATACACTAGAG ATACCATGAG GTGATAACAC CCAGATGCG GAGCTAACCC CCTTTTTTGCA ACCAAACGAC CCAGCATGCC CCAGCATGCC CCAGCACCAC CCACCAACCAC CCACCAACCAC CCACCAACCAC CCACCA			, acompound of the part of the	TOTAL COLOR	IAITAILLUG	THITOMACOCC	7020
CCAGTCACAG AAAAGCATCT TACGGATGGC ATGACAGTAA GAGAATTAC CAACGATCG AGACCGAAG 71 ATAACCATGA GTGATAACAC TGCGGCCAAC TTACTTCTGA CAACGATCGG AGGACCGAAG 72 GAGCTAACCG CTTTTTTGCA CAACATGGGG GATCATGTAA CTCGCCTTGA TCGTTGGGAA 72 CCGGAGCTGA ATGAAGCCAT ACCAAACGAC GAGCGTGACA CCACGATGCC TGTAGCAACA 72 GCAACAACGAT TGCGGCAAACT ATTAACTGGC GAACTACTTA CTCTAGCTTC CCGGCAACAA 73 GCAACAACGAT GGATGAGGC GGATAAAGTT GCAGGACCAC TTCTGCGCTC GGCCCTTCCG 73 GCAGCACTGG GGCCAGATGG TAAACTCTGGA GCCGGTGAGC GTGGGTCTCG GGCCCTTCCG 73 GCAGCAACTA TGGATGAACG AAATAGACAG ATCGCTGAGA TAGGTGCCTC ACTGATTAAG 74 CATTGGTAAC TGGATGAACG AAATAGACAG ATCGCTGAGA TAGGTGCCTC ACTGATTAAG 74 TAACGTGAGT TTTCCTCCA CTGAGCGTCA GACCCCGTAG AAAAGAACCAC CAAAATCCCT 74 TAACGTGAGT TTTTCTCCCA CTGAGCGTCA GACCCCGTAG AAAAGAACCA ACCGCTACCA 74 TGGAGTCCTT TTTTTCTCGC CGTAATCTGC TGCTTGCAAA ACCTGCTTC 74 AGCAGAGCGC AGATACCAAA TACTGTTCTT CTAGTGTAGC CAAAAAAAACC ACCGCTACCA 74 AGCAGAGCGC AGATACCAAA TACTGTTCTT CTAGTGTAGC CGAACTACTT 74 AGCAGGGGG AGAACCAAA TACTGTTCTT CTAGTGTAGC CGACCACTTC 74 AGCAGGGGG AGATACCAAA TACTGTTCTT CTAGTGTAGC CGACCACTTC 74 AGCAGGGGG AGATACCAAA TACTGTTCTT CTAGTGTAGC CGACCACTTC 75 AAGAACTCTG TAGCACCGCC TACATACCTC GCTCTGCTAA TCCTGTTAGC ACCGGATAAG 75 GCCGGTGGCG ATAAGTCGT TCTTACCGGG TTGGACTACA ACCGGATAAG 75 ACACCGAAC TGAGATACCT ACAGCGTCA ACAGCTTCAA ACCGGATAAG 75 ACACCGAAC TGAGATACCT ACAGCGTCAA ACAGCTTCAA ACCGGATAAG 75 AGAAAGGCGG ACAGCTTCG GGGGGGTTCG TGCACACAC CCACCTTCGACT 75 AGAAAGGCGG ACAGCTTCG GGGGGGTTCG TGCACACAC CAAAACACC CCACCACTTC 75 ACACCGAAC TGGAGTACC ACAGCTTCAA ACCGGATAAG 75 ACACCGAAC TTCTTACCGG TTGCACACAC CAAAAAAAAAA		~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~	י הההההאתימני	TATITUAGA	NIGWC I I GG I	10110 1110 1	7080
ATAACCATGA GTGATAACAC TGGGGCCAAC TTACTTCTGA CAACGATGGG AGGGCTAACCG CTTTTTTGCA CAACATGGGG GATCATGTAA CTCGCCTTGA TCGTTGGGAA 72 CCGGGAGCTGA ATGAAGCCAT ACCAAACGAC GACCGTGACA CCACGATGCC TGTAGCAATG 72 CCGGAACAACGT TGCGCAACACAC ACCAAACGAC GACCGTGACA CTCTAGCTTC CCGGCAACAA 73 CCAACAACGT TGCGCAACACAC TTTAATAGACT GGATGAGGG GGATAAAGTT GCAGGACCAC TTCTGCGCTC GGCCCTTCCG 73 CCGGCAACAA 74 CCAACAACGT TTAATTGCTGA TAAATCTGGA GCCGGTGAGC GTGGGTCTCG CGGTATCATT 74 CCGGCAACCAC TTCTGCGCTC CGGTATCATT 74 CCGGCAACCAC TTCTGCGCTC CGGTATCATT 74 CCGGCAACCAC TGGATGACG ACTGATCACAC ACCCCGTAGAC ACCCCCTAGAC ACCCCGTAGAC ACCCCCTACACAC TTCTTCTCCAC CTGAGCGTCA GACCCCCGTAG AAAAGAACCAC CAAAATCCCT 75 CCGGTGGTTT TTTTCTGCG CGTAATCTGC TGCTTGCAAA ACCGCTACCA 75 CCGGTGGTTT TTTTCTGCG CGTAATCTGC TGCTTGCAAA ACCGCTACCA 75 CCGGTGGGC ACAGAACACAC ACCGCTACCA 75 CCGGTGGGC ACAGAACCACAC TCCTGTTCTT CTAGTGTAGC CCACCACTTC 75 CCGGTGGGCG ACAGACCCCC TACATACCTC GCTCTGCTAA TCCTGTTACC AGTGGCTGC 75 CCGGGGGGGC ACAGATACCTA ACAGCGTCAA ACCGGTTAAG ACCGGATAAG ACCGGTTACAC ACCGGATAAG ACCGGATAAG ACCGGTTACAC ACCGGTTACAC ACCGGATAAG ACCGGTTACAC ACCGGATAAG ACCGGTTACAC ACCGGATAAG ACAGACAAG ACCGGATAAG ACCGGATAAG ACCGGATAAG ACCGGATAAG ACCGGATAAG A			n marrenamere	ATTAL ALTIA	I GWGWWTTWIG	CMOICCIGG	7140
GAGCTAACCG CTTTTTTGCA CAACATGGGG GATCATGTAA CTCGCCTTGA CCGGAGCTGA ATGAAGCCAT ACCAAACGAC GAGCGTGACA CCACGATGCC TGTAGCAATG GCAACAACGT TGCGCAAACT ATTAACTGGC GAACTACTTA CTCTAGCTTC CCGGCAACAA TTAATAGACT GGATGGAGGC GGATAAAGTT GCAGGACCAC TTCTGCGCTC GGCCCTTCCG GCTGGCTGGT TTATTGCTGA TAAATCTGGA GCCGGTGAGC GTGGGTCTCG GGCCCTTCCG GCAGCACTAG GGCCAGATGG TAAGCCCTCC CGTATCGTAG TTATCTACAC GACGGGAGT CAGGCAACTA TGGATGAACG AAATAGACAG ATCGCTGAGA TAGGTGCCTC ACTGATTAAG CATTGGTAAC TGTCAGACCA AGTTTACTCA TATATACTTT AGGTTGCTC AAAAACCTCT TTTTAATTTA AAAGGATCTA GGTGAAGATC CTTTTTGATA ATCTCATGAC CAAAATCCCT TGAGATCCTT TTTTCTGCG CGTAATCTGC TGCTTGCAAA CAAAAAAACC ACCGCTACCA TGCGGTGGTT GTTTGCCGGA TCAAGAGCTA CCAACTCTTT TCCGAAGGGT AACTGGCTTC AGCAGAGCGC AGATACCAAA TACTGTTCTT CTAGTGTAGC CGTAGTTAGG CCACCACTTC GCCGTGGCG ATAAGTCGTG TCTTACCGGG TCGACCACAC TCACGAACACAC ACCGCTACCA TACAACCCGAAC TGCACACAC TCTGCTCTC TTGGACTCAA CACGCATAGGT ACCGGATAAG GCCAGTGGCG ATAAGTCGTG TCTTACCGGG TTGGACTCAA GACGATAGTT ACCGGATAAG GCCAGTGGCG ATAAGTCGTG TCTTACCGGG TTGGACTCAA GACGATAGTT ACCGGATAAG GCCAGTGGCG ATAAGTCGTG TCTTACCGGG TTGGACTCAA GACGATAGTT ACCGGATAAG GCCAGTGGCG ATAAGTCGTG TCTTACCGGG TTGGACACAC CACGACTTC TTGGACACACAC CGCCACACTTC TACACCGAAC TGGACTACCT ACCAGCGTGAAC CAGGAGAACAC CAGGAGAACAC ACCGCTACAC TACACCGAAC TGGACTACCT ACCACACACAC CAGGCTTGGA CAGGAGAGCC CACCACTTC TACACCGAAC TGGACTACCT ACCACACACAC CAGGCTTGGA CAGGAGAGCC CACCACACTT ACACCGAAC TGGACTACCT ACCACCGC TTATGAAAAA CCCCGAACGAC CTCCCGAAGGG AGAAAGGCGG ACAGGTATCC ACGGCTGGAA CAGGGTCGAA CAGGAGAGCC CACCACACTT ACACCGAAC TCCCGAAGGG CAGGAGACC CTCCCGAAGGG AGAAAGGCCG ACAGGTATCC ACGGCTGAAC CTCCTGCCA CCTCTGACTT AGACCGGAT TTTTTTTTATATACTTTAT AGTCCTGCG GTTTCGCCA CCTCTGACTT AGACCGGAT TTTTTTTTTATATACTTTAT AGTCCTGCG GGTTCCGCA CCTCTGACT AGACCGCAC TTTTTTTTTATATACTTTATATACTTTTATATGTTC TTTTTTTTATATATACTTTATATACTTTTATATACTTTTATATACTTTTATATACTTTTATATACTTTTATATACTTTTATATACTTTTATATACTTTTATATACTTTTATATACTTTTATATACTTTATATACTTTTATATACTTTTATATACTTTTATATACTTTTATATACTTTTATATACTTTTATATACTTTTATATACTTTTATATACTTTTATATACTTTTATATACTTTTATATACTTTTATATACTTT			, <i>madadar</i> raat	' ''''' A C '' I "I 'U ' I U - A	CHACGHICGC	, ,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,	7200
CCGGAGCTGA ATGAAGCCAT ACCAAACGAC GACCTTCACA CCACACTGC CCAACAACGT TGCGCAAACT ATTAACTGC GAACTACTTA CTCTAGCTTC CCGGCAACAA TTAATAGACT TTAATAGACT GGATGAGGC GGATAAAGTT GCAGGACCAC TTCTGCGCTC GGCCCTTCCG GCTGGCTGGT TTATTGCTGA TAAATCTGGA CAGGCAACTA TGGATGAACG TTAATCTGA CAGGCAACTA TGGATGAACG TAAATCTCGC GCCGTTAGCT CCGGTATCATT AAAACTTCAT TAACTGTAAC TTTTAATTTA AAAGGATCTA TGGATGACCA AGTTTACTCA TATATACTTT AAAAGATCAT TTTTTAATTTA AAAGGATCTA TTTTTCTCA CGGATACTCC CGTATCTTT TTTTCTCAGAC AAAAACACC CTTTTTTAATTTA AAAGGATCTA TTTTTCCAGACCA TACATCTCC TTTTCCAGACCA AGAACCCCT TTTTTCCAGACCA TACATCTCC TTTCCAGACCA TACATCCTC TTTCCAGACCA TACATCCTC TTTCCAGACCA TACATCCTC TTTTCCAGACCA TACATCCTC TTTGCCAGAC TACATCCTC TTTTCCAGACCA TACATCCTC TTTGCCAGAC TCCTAGACACC TTCCAGACCAC TTCCACACCC TTCCACACC TTCCACACCC TTCCACACC TTCCACACCC TTCCAC			* ~ X X ~ X ~ X ~ CCCCC	: CATCATGTA	I CICCCTIC	7 10011000	7260
GCAACAACGT TGCGCAAACT ATTAACTGC GAACTACTTA TCTGTGCGTC GGCCCTTCCG TTAATAGACT GGATGAGGC GGATAAAGTT GCAGGACCAC TTCTGCGCTC GGCCCTTCCG TTAATAGACT TTATTGCTGA TAAATCTGGA GCCGGTGAGC GTGGGTCTCG CGGTATCATT 74 GCAGCACTGG GGCCAGATGG TAAGCCCCTCC CGTATCGTAG TTATCTCACAC GACGGGAGT 75 CAGGCAACTA TGGATGAACG AAATAGACACA ATCGCTGAGA TAGGTGCCTC ACTGATTAAG 75 CATTGGTAAC TGTCAGACCA AGTTTACTCA TATATACTTT AGATTGATT AAAACTTCAT 77 TAACGTGAGT TTTCGTTCCA CTGAGCGTCA GACCCCGTAG AAAAGAACCA ACGGTTCCT 77 TGAGATCCTT TTTTTCTGCG CGTAATCTGC TGCTTGCAAA AAAAAAAACC ACCGCTACCA 77 AGCAGAGCGC AGATACCAAA TACTGTTCTT CTAGTGTAGC CCAACATCTC 77 AGCAGAGCGC AGATACCAAA TACTGTTCTT CTAGTGTAGC CCACCACTTC 77 GCCAGTGGCG ATAAGTCGT TCTTACCGGG GCCACCACT 77 GCCAGTGGCG ATAAGTCGT TCTTACCGGG TGCACAAAAAAAACC ACCGGATAAG 86 GCCAGTGGCG ACAGGTAACCT TCTTACCGGG TGCACAACACC CCACCACTTC 77 AGCACCGAAC TACATACCTC TTGGACTCAA GACGATAGTT ACCGGATAAG 86 GCCAGTGGCG ACAGGTAACCT ACAGCGCCCACACCC CCACCACTTC 77 AGAAACGCGG ACAGGTAACCT ACAGCGCGC CACGACGCC TCCCGAAGGG GCGAACGACC CCACCACTTC 77 GCCAGTGGCG ACAGGTAACCT ACAGCGCGC CTATGAAAA CCCCGAACGAC CCACCACTTC 77 AGAAAGGCGG ACAGGTAACCT ACAGCGCGC CTATGAAAA CCCCGAACGAC CCACCACTTC 77 AGCACCGGA CACGACGCC TCCCGAAGGG ACAGCGCC CACGAACGAC CCTCCGAAGGG ACAGAACGCC CTTCCAAGGG ACAGACGCC CACGAACGAC CCTCCGAAGGG ACAGAACGCC CTTCCAAGGG ACAGCGCC CACGAAGGAC CCTCCGAAGGG ACAGCGCC CTCCGAACGAC CCTCCGAACGAC CCTCCCAACGC CCTCCGAACGAC CCTCCCAACGC CCTCCGAACGAC CCTCCCAACGC CCTCCGAACGAC CCTCCCAACGC CCTCCCAACGC CCTCCGAACGC CACCACCTT TCCCGAAGGG ACAGCCCC CTCCCAACGC CCTCCGAACGAC CCTCCCAACGC CCTCCCAACGC CCTCCGAACGAC CCTCCCAACGC CCCACCACCCC CCCCCACCACCC CCTCCCAACGC CCTCCCAACCACC CCTCCCAACCACC CCTCCCAACCACC CCTCCCAACCACC CCTCCCAACCACCC CCCCCACCACCC CCCCACCACCC CCCCACCA			<b>** **********</b>	' CALCETTACE	1 CCMCGM1GC	. 10111000	7320
TTAATAGACT GCATGAGAGC GCATAAAGTT GCAGGACCAC GCTGGCTGGT TTATTGCTGA TAAATCTGGA GCCGGTGAGC GGCAGCACTG GGCCAGATGG TAAGCCCCTCC CGTATCGTAG TAGGTGCCTC CAGGCAACTA TGGATGAACG AATAGACCA ATCGCTGAGA TAGGTGCCTC TTTTAATTTA AAAGGATCTA TAGGTGAGCA TTTTCTTCCA TTTTAATTTA AAAGGATCTA TTTTCTCCA CGTAATCTCC TTTTTCTCCA CGTAATCTCC TTTTTCTCCA TAGCACCACAC TTTTTCTCCA TACAGACCA AGATACCACA TACTCTTCT TTTCCGGAGC TACAGACCTA TACTCTTCT TTTCCGAAGC AAAAAAAACC ACCGCTACA TACTCTTCT TTCCGAAGCT AGATACCAC TACAGACCTA TACTCTTCT TTCCGAAGCT TACAGACCTA TACTCTTCT TTCCGAAGCT TACAGACCTA TACTCTCT TTTCCGAAG TACACCACAC TTTTCCGAAC TACATACCTC TTTTACCGG TTCTTACCGGG TTCCAGAGCC TTCCAGAGC TTCCAGAGCC TTCCAGAGC TTCCAGAGCC TTCCAGAGCC TTCCAGAGCC TTCCAGGGG TTCCTAGCGCC TTCCAGGGG TTCCTAGCCCC TTCCAGGGG TTCCTCAGGC TTTTTGTGATC TTTTGTGATC TTTTGTGATC TTTTCTCAGGC TTCCAGAGCC TTTTTGTGATC TTTTTTTTTT			ייין די אירוידיא אירידיירייניי	' /2ΔΔ('''ΔL 1517	4 1.11.1200	, ,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,	7380
GCTGGCTGGT TTATTGCTGA TAAATCTCGA GCAGCACTGG GGCCAGATGG TAAGCCCTCC CGTATCGTAG TAGGTGCCTC CAGGCAACTA TGGATGAACG AAATAGACAG ATCGCTGAGA TAGGTGCCTC ACTGATTAAC TTTTAATTTA AAAGGATCTA GGTGAAGATC TTTTCTCCA TTTTCTCCC CGTATCTTA TACCGTGAC TTTTCTCCC CGTATCTTT AAAAGATCAT TATATACTTT AAAAGATCAT TATATCTCA TACCGTGAGT TTTTCTTCCA CGTAATCTCC TTTTTCTCCC CGTAATCTCC TTTCCAGACC TCAAGACCTA TCCTGCCAAA AAAAAAACC ACCGCTACCA TTTCGCGGA TCAAGACCTA TCCTGCTACA AAAAAAAACC ACCGCTACCA TACCGCTCC TACAACCTTT TTCCGAAAGT AACTGCTTC TTTTCCCGAAAGT TACCTGCTAC TCCTGGCTACA TACCTGCTC TACAACCTC TTTTACCGG TACATACCTC TTTTACCGG TTCTACCGGG TTCTACCGGG TTCTACCGGG TTCTACCGGG TTCTACCGGG TTCTACCGGG TTCTACCGGG TTCTACCGGG TTCTACCGGG TTCCAAGACCC TTCCAAGACC TTTTTCTTTTT TTTTTTTTTT			מיים אות אידות אידות	' (-( A(-(-A(-(-A(	110100001	, 00000	7440
CAGGCAACTA TGGATGAACG AATTAGACAG ATCCTTGAAA TAGATTGATT AAAACTTCAT TGTTGAGACCA AGTTTACTCA TATATACTTT AGATTGATT AAAACTTCAT TTTTTAATTTA AAAGGATCTA GGTGAAGATC CTTTTTGATA ATCTCATGAC CAAAAATCACT TAACGTGAGT TTTCGTTCCA CTGAGCGTCA GACCCCGTAG AAAAGAACAC ACCGCTACCA TGAGAGCTA CAAAAAAAACC ACCGCTACCA TGCTGAGAGCTA CAAAAAAAACC ACCGCTACCA TAGAGAGCTA CCAACTCTTT TTCCGAAAGT AACTGGCTTC TAGCAGAGCTA TACTGTTCTT CTAGTGTAGC CGTAGTTAGG CCACCACTTC TAGCAGCGC TACATACCTC GCTCTGCTAA GACGATAGTT ACCGGATAAG GCGCAGCGGC TACATACCTC GCTCTGCTAA GACGATAGTT ACCGGATAAG GCGCAGCGGC CGGGCTGAAC GGGGGGTTCG TGCACAACAGC CCAGCTTGGA GCGAACGACC GGGGGGTTCG TGCACACAGC CCAGCTTGGA GCGAACGACC GCACACTC TCCCGAAGGG ACAGCTACCA ACAGCGTGAA AGAAAGGCGG ACAGGTATCC GGTAGCTGAA AGGGCGGAACGACC GTACCTTTATA AGTCCTGCGA CAGGAGAGAC CACCGAGGGA GAGCGCTCTTTTTTTTTT			* 10 7 7 7 7 10 11 11 11 11 11 11 11 11 11 11 11 11	1 (-( (-(-) ) ) A	" GIRRATOR	,	7500
CATTGGTAAC TGTCAGACCA AGTTTACTCA TATATACTT AGAITATACTT TGTTAATTTA AAAGGATCTA GGTGAAGATC CTTTTTGATA ATCTCATGAC CAAAATCCCT TTTTAATTTA AAAGGATCTA GGTGAAGATC CTTTTTGATA ATCTCATGAC CAAAAATCA AGGATCTTC TAACGGTGAGT TTTTGTTCCAG CGTAATCTGC TGCTTGCAAA CAAAAAAAACC ACCGCTACCA TGCGGTGGTTT GTTTGCCGGA TCAAGAGCTA CCAACTCTTT TTCCGGAAGGT AACTGGCTTC TAGCAGAGCTA TACTGTTCTT CTAGTGTAGC CGTAGTTAAC CCACCACTTC TAGCAGCGC TACATACCTC GCTCTGCTAA TCCTGTTACC AGTGGCTGCT TCCTACCGG TTGGACTCAA GACGATAGTT ACCGGATAAG GCGCAGCGC CCACCACTC TTGCACACAGC CCAGCTTGGA GCGAACGACC GGGCGAGCGC TACACACAGC CCACCACTTC TCCCGAAGGG ACAGCGCAACGC CCAGCTTGGA GCGCACGACC CCACCACTC TCCCGAAGGG ACAGCGCAACGC CTCCGAACGAC CCACCACTC TCCCGAAGGG GAAAAGGCCG GAAACGCCTG GTAACTTTATA AGTCCTGTCG GGTTTCGCCA CCTCTGACTT GAGCGCCACCTT TTTTTGTGATG CTCCTCAGGG GGGCGCGCACGCT TACCCAAGGCT TTTTTGTGATG CTCCTCAGGG GGGCGCGCAT GGATATGTTC TGCCAAGGGT ACCCAACGCC CTCCTGACTT TTTTTGTGATG CTCCTCACGG GGGCGCGCAT GGATATGTTC TGCCAAGGGT ACCCAACGCC CTCCTCACACACC CTCTCACACACCC CTCCTCACACACC CTCCTCACACACC CTCCTCACACACCC CTCCTCACACACC CTCCTCACACACCC CTCCACACACCC CTCCTCACACACCC CTCCTCACACACCC CTCCTCACACCC CTCCTCACACACCC CTCCTCACACACCC CTCCCACACCC CTCCTCACACACCC CTCCTCACACCCCCCCC	GCAGCACTG	G GGCCAGATG	G TAAGCCCTCC	CGTATUGIA	A TAGGTGCCT	ACTGATTAAG	7560
TTTTAATTTA AAAGGATCTA GGTGAAGATC CPPTTGATA ATCHARGAC AGGATCTTCT TAACGTGAGT TTTCGTTCCA CTGAGCGTCA GACCCCGTAG AAAAGAATCA AGGATCTTCT TAACGTGAGT TTTTTTCTGCG CGTAATCTGC TGCTTGCAAA CAAAAAAACC ACCGCTACCA TGCGGTGGTTT GTTTGCCGGA TCAAGAGCTA CCAACTCTTT TTCCGGAAGGT AACTGGCTTC TAGCAGCGCA TCAAGAGCTA CCAACTCTTT TTCCGGAAGGT AACTGGCTTC TAGCAGCGCC TACATACCTC GCTCTGCTAA TCCTGTTACC AGTGGCTGCT TACATACCTC GCTCTGCTAA TCCTGTTACC AGTGGCTGCT TCCTACCGGG GCGAACGACC CCAGCTTGGA GCGCACGACC CCAGCAGAGG GCGCACGACC CCAGCAGAGG GCGCACGACC CTCCGAAGGG ACAGCACC GGTAAGCGC AGGGGAGAGAC CAGGAGAGGC CACGAGGGAG GCGCACGCT TCCCGAAGGG GAAAAGGCGG GAAACGCCTG GTAACTTTATA AGTCCTGTCG GGTTTCGCCA CCTCTGACTT GAGCGCTCACG GCGCGCGCC TTTTTGGAAGAAA CGCCAGCAAC CTCCTCAGATG GGGCGCGCC TATGGAAAAA CGCCAACAAC CTCCTCAGATGC CTCCTCAGGG GGGCGCGCC TATGGAAAAA CGCCAACAAC CTCCTCAGATGC CTCCTCAGGT GGGCGCGCC TATGGAAAAA CGCCAACAAC CTCCTCAGATGC CTCCTCAGATGC CTCCTCACGT GGGTATGTTC TGCCAAGGGT CTCCTCAGATGC CTCCTCACGT GGGTATGTTC TGCCAAGGGT CTCCTCAGATGC CTCCTCACGT GGGTATGTTC TGCCAAGGGT CTCCTACACTT GCCAAGGGT CTCCTCACGT GGGTATGTTC TGCCAAGGGT CTCCTCACGT GGATATGTTC TGCCAAGGGT CTCCTCACGT GGGTATGTTC TGCCAAGGGT CTCCTCACGT GGATATGTTC TGCCAAGGGT CTCCTCACGT GGATATGTTC TGCCAAGGGT CTCCTCACGT GGATATGTTC TGCCAAGGGT CTCCTCACGT GGATATGTTC TGCCAAGGGT CTCCTACACACC CTCCTCACACACC CTCTCACACACC CTCCTCACACACC CTCCTCACACACC CTCCTCACACACC CTCCTCACACACA	CAGGCAACT	A TGGATGAAC	G AAATAGACAG	ATCGCIGAG	T ACATTCATT	T AAAACTTCAT	7620
TAACGTGAGT TTTCGTTCCA CTGAGCGTCA GACCCCGTAA AAAAAAAACC ACCGCTACCA TGAGATCCTT TTTTTCTGCG CGTAATCTGC TGCTTGCAAA CAAAAAAAACC ACCGCTACCA TGAGAGCTA CCAACTCTTT TTCCGAAAGGT AACTGGCTTC TAGCAGAGCTA CCAACTCTTT TTCCGAAAGGT AACTGGCTTC TAGCAGAGCTA TACTGTTCTT CTAGTGTAGC CGTAGTTAGC CCACCACTTC TAGCAGCACA TACTGTTCTT CTAGTGTAGC CGTAGTTACC AGTGGCTGCT TACATACCTC GCTCTGCTAA TCCTGTTACC AGTGGCTGCT TCCTACCGGG TCCTACCAGC CCAGCATGGA GCGCAACGACC GGGCAGCGC TCCTCAGAGGG GCGAACGACC CCAGCAGGGA GCGCACGACC CACGAAGGG ACAGGATACCT ACACGCGGAAC AGGGTCGGAA AGGGCCGAACGACC GGAAAAGGCGG ACAGGTATCC GGTAACCTGA AGGCTCGGAA CGCCACGCAC CCTCTGACTT GAGCCGCCACCAC CTCCTCAGGG GAAACGCCTG GTATCTTTAT AGTCCTGTCG GGTTTCGCCA CCTCTGACTT GAGCCGCCACCACC CTCTGACTT CTCCTCAGGG GGCCGCGCCC TATGGAAAAA CGCCAACAAC CTCCTCAGATGC CTCCTCAGGT GGGCGCGCAT GGATATGTTC TGCCAAGGGT	CATTGGTAA	C TGTCAGACC	A AGTTTACTC	- COMMUNICAT	A ATCTCATGA	CAAAATCCCT	7680
TGAGATCCTT TTTTTCTGCG CGTAATCTGC TGCTTGCAAA CAAAAGACTC TGCGGTGGTTT GTTTGCGGA TCAAGAGCTA CCAACTCTTT TTCCGGAAGGT AACTGGCTTC TAGCAGCGAAA TACTGTTCTT CTAGTGTAGC CGTAGTTAGG CCACCACTTC TAGCAGCGC TACATACCTC GCTCTGCTAA TCCTGTTACC AGTGGCTGCT TACATACCTC GCTCTGCTAA TCCTGTTACC AGTGGCTGCT TCCTACCGGG TCCTACCACAGC CCAGCATAGT ACCGGATAAG GCGCAACGACC GGGCAGCGC TCCCGAAGGG GCGAACGACC CCAGCTTGGA GCGCACGACC CCAGCAGGGA GCGCAACGACC CAGCAGGGAAAAAAAA	TTTTAATTI	A AAAGGATCT	A GGTGAAGAT	Z CACCCCGTA	G AAAAGATCA	A AGGATCTTCT	7740
GCGGTGGTTT GTTTGCCGGA TCAAGAGCTA CCAACTCTTT TICCGAAGAGT CCACCACTTC TAGCAGCGC AGATACCAAA TACTGTTCTT CTAGTGTAGC CGTAGTTAGG CCACCACTTC TAGCAGCGC TACATACCTC GCTCTGCTAA TCCTGGTTACC AGTGGCTGCT TCCAGTGGGG TCCTTACCGGG TTGGACTCAA GACGATAGTT ACCGGATAAG GCGCAGCGGC CCAGCTTGGA GCGAACGACC GCGCAGCGGC CCAGCTTGGA GCGCACGACC CCAGCTTGGA GCGCAACGACC ACAGAAAGGCGG ACAGGTATCC GGTAAGCGG AGGGTCGGA CAGGAGAGGG CACGAGGGAG GAAAGGCGG GAAACGCCTG GTATCTTTAT AGTCCTGTCG GGTTTCGCCAA CGCCAGCAAC GAGCGTCGAT TTTTGTGATG CTCGTCAGGG GGGCGGAGCC TATGGAAAAA CGCCAGCAAC GTCCTCAGAGTG GGGCGGAGCC TATGGAAAAA CGCCAGCAAC GCCCAGCGAAC GCCCAGCGAAC GCCCAGCGAAC GCCCAGCGAAC GCCCAGCGAAC GCCCAGCGAAC GGGCGCGCACC TTTTGGGATG CTCGACAATGG CGGCCGGAT GGATATGTTC TGCCAAGGGT	TAACGTGAG	T TTTCGTTCC	A CTGAGCGIC	C TCCTTCCAA	A CAAAAAAAC	C ACCGCTACCA	7800
AGCAGAGCGC AGATACCAAA TACTGTTCTT CTAGTGTAGC CGTATTAGC AGTGGCTGCT 7  AAGAACTCTG TAGCACCGCC TACATACCTC GCTCTGCTAA TCCTGTTACC AGTGGCTGCT 7  GCCAGTGGCG ATAAGTCGTG TCTTACCGGG TTGGACTCAA GACGATAGTT ACCGGATAAG 8  GCGCAGCGGT CGGGCTGAAC GGGGGGTTCG TGCACACAGC CCAGCTTGGA GCGAACGACC 8  TACACCGAAC TGAGATACCT ACAGCGTGAG CTAGAGAAAA GCGCCACGCT TCCCGAAGGG AGAAAAGGCGG ACAGGTATCC GGTAACGGC AGGGTCGGAA CGCTCTGACTT 8  GAGCGTCGAT TTTTGTGATG CTCGAGATGG CGGCGGAGCC TATGGAAAAA CGCCAGCAAC 6  GAGCGTCGAT TTTTGTGATG CTCGAGATGG CGGCGCGAT GGATATGTTC TGCCAAGGGT 8	TGAGATCCT	T TTTTTCTGC	G CGTAATCIG	Z CCAACTCOT	T TTCCGAAGG	T AACTGGCTTC	7860
AAGAACTCTG TAGCACCGCC TACATACCTC GCTCTGCTAA TCCTGTAACTCTAACGACGACGACGACGACGACGACGACGACGACGACGACG	GCGGTGGTT	T GTTTGCCGC	A TCAAGAGCI	m cmacremas	C CGTAGTTAG	G CCACCACTTC	7920
GCCAGTGGCG ATAAGTCGTG TCTTACCGGG TTGGACTCAA GACGATAGTT GCGAACGACC GGGCAGCGC CCAGCTTGGA GCGAACGACC GCGCAGCGGT TCCCGAAGGG GCGCAGCGAAC TGAGATACCT ACAGCGTGAG CTATGAGAAAA GCGCCACGCT TCCCGAAGGG AGAAAAGGCGG ACAGGTATCC GGTAACCTTATA AGTCCTGTCG GGTTTCGCCA CCTCTGACTT GAGCGTCGAT TTTTGTGATG CTCGTCAGGG GGGCGGAGCC TATGGAAAAA CGCCAGCAAC GAGCGTCGAT TTTTGTGATG CTCGTCAGGG GGGCGGAGCC TATGGAAAAA CGCCAGCAAC GAGCGTCGAT TTTTGTGATG CTCGTCAGGG GGCGGAGCC TATGGAAAAA CGCCAAGGGT GAGCGTCGACAC GTCGACAATGG CTCGACAATGG CGGACGCGAT GGATATGTTC TGCCAAGGGT	AGCAGAGC	C AGATACCA	A TACTGIICI	C CCTCTCCTA	A TCCTGTTAC	C AGTGGCTGCT	7980
GCGCAGCGGT CGGGCTGAAC GGGGGGTTCG TGCACAGC CCACGCTTGCT TCCCGAAGGG TACACCGAAC TGAGATACCT ACAGCGTGAG CTATGAGAAA GCGCCACGCT TCCCGAAGGG AGAAAGGCGG ACAGGTATCC GGTAAGCGGC AGGGTCGGAA CAGGAGAGCG CACGAGGGAG CTTCCAGGGG GAAACGCCTG GTATCTTTAT AGTCCTGTCG GGTTTCGCCA CCTCTGACTT GAGCGTCGAT TTTTTGTGATG CTCGTCAGGG GGGCGGAGCC TATGGAAAAA CGCCAGCAAC GAGCGTCGACT CTCGTCAGATGG CGGACGCGAT GGATATGTTC TGCCAAGGGT	AAGAACTC	rg Tagcaccgo	TACATACCI	C TOTOLOGIA	A GACGATAGT	T ACCGGATAAG	8040
TACACCGAAC TGAGATACCT ACAGCGTGAG CTATGAGAAAA GCGCCACGGC CACGAGGGAG AGAAAAGGCGG ACAGGTATCC GGTAAGCGGC AGGGTCGGAA CAGGAGAGCG CACGAGGGAG CTTCCAGGGG GAAACGCCTG GTATCTTTAT AGTCCTGTCG GGTTTCGCCA CCTCTGACTT GAGCGTCGAT TTTTGTGATG CTCGTCAGGG GGGCGGAGCC TATGGAAAAA CGCCAGCAAC GAGCGTCGAT TTTTGTGATG CTCGAGATGG CGGACGCGAT GGATATGTTC TGCCAAGGGT	GCCAGTGG	JG ATAAGTCG	C CCCCCCTTC	C TCCACACAC	C CCAGCTTGG	A GCGAACGACC	8100
AGAAAGGCGG ACAGGTATCC GGTAAGCGGC AGGGTCGGAA CAGGAGAGCG CACCAGGACTT  CTTCCAGGGG GAAACGCCTG GTATCTTTAT AGTCCTGTCG GGTTTCGCCA CCTCTGACTT  GAGCGTCGAT TTTTGTGATG CTCGACATCG CGGACGCGAT GGATATGTTC TGCCAAGGGT	GCGCAGCG	ST CGGGCTGA	AC GGGGGGIIC	C CTATGAGAE	A GCGCCACGC	T TCCCGAAGG	8160
CTTCCAGGGG GAAACGCCTG GTATCTTTAT AGTCCTGLCG GGTTTCGCCA CGCCAGCAAC  GAGCGTCGAT TTTTGTGATG CTCGACATCG CGGACGCGAT GGATATGTTC TGCCAAGGGT				Y A(2(2(2))) (3)	VA LAUGAGAG	.0	
GAGCGTCGAT TTTTGTGATG CTCGTCAGGG GGGCGGAGCC TATGGAAAAA COOCHE				17 At 11 115 11	(56) 11 (60)	,,,, ,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,	
TOTAL			**C CTCCTCA(()	JEJMEDEL Lelevie	" TWIGOWWW	B# CCCO.	
TGGTTTGCGC ATTCACAGTT CTCCGCAAGA ATTGATTGGC TCCAATTCTT GGAGTGGTGA			TO CTCCACATO	C (CALLE	I GOWINIGH		_
TGGTTTGCGC ATTCACAGII GIGGGGIIIII III	GGATGCGC	CG CGTGCGGC	TT CTCCGCAAC	A ATTGATTG	SC TCCAATTC	T GGAGTGGTG	A 8460
	TGGTTTGC	GC ATTUACAG	11 Oldcoons				

WO 00/15819 PCT/US99/20675

### 28/29 :

ATCCGTTAGC GAGGTGCCGC	CGGCTTCCAT	TCAGGTCGAG	GTGGCCCGGC	TCCATGCACC	8520
GCGACGCAAC GCGGGGAGGC	AGACAAGGTA	TAGGGCGGCG	CCTACAATCC	ATGCCAACCC	8580
GTTCCATGTG CTCGCCGAGG					8640
AGTTAGGCTG GTAAGAGCCG					8700
CTGCCTGGAC AGCATGGCCT					8760
CATAATGGGG AAGGCCATCC					8820
CAAAAAGCC TCCTCACTAC	TTCTGGAATA	GCTCAGAGGC	CGAGGCGGCC	TCGGCCTCTG	8880
CATAAATAAA AAAAATTAGT	CAGCCATG 8	3908			

Fig. 10D

\*\*

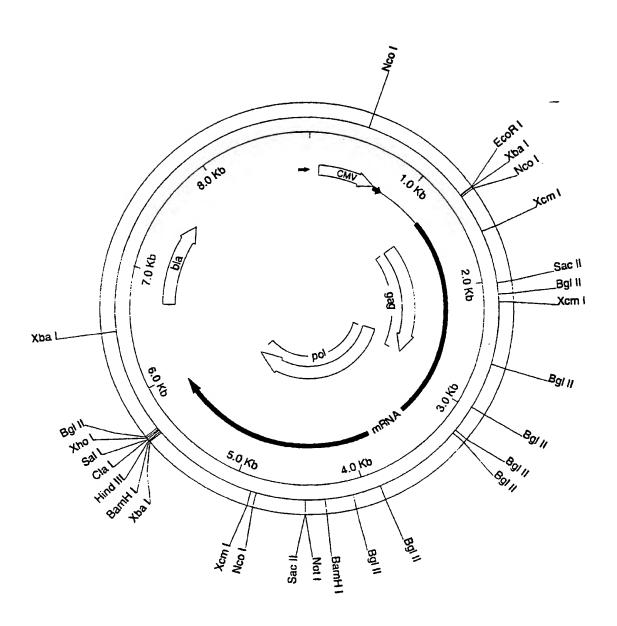


Fig. 11

# INTERNATIONAL SEARCH REPORT

r national Application No

		101/03 99/20075
A. CLASSI IPC 7	FICATION OF SUBJECT MATTER C12N15/86 C12N5/10 C12N7/0	4 C12N15/49 C07K14/16
According to	o International Patent Classification (IPC) or to both national classifi	cation and IPC
	SEARCHED	
IPC 7	ocumentation searched (classification system followed by classification classification system followed by cl	ion symbols)
Documenta	tion searched other than minimum documentation to the extent that	such documents are included in the fields searched
Electronic d	lata base consulted during the international search (name of data b	ase and, where practical, search terms used)
C. DOCUM	ENTS CONSIDERED TO BE RELEVANT	
Category °	Citation of document, with indication, where appropriate, of the re	olevant passages Relevant to claim No.
A	NALDINI L ET AL: "IN VIVO GENE AND STABLE TRANSDUCTION OF NONDI CELLS BY A LENTIVIRAL VECTOR" SCIENCE, US, AMERICAN ASSOCIATION ADVANCEMENT OF SCIENCE, vol. 272, no. 5259, 12 April 1996 (1996-04-12), page XP000583652 ISSN: 0036-8075 cited in the application the whole document	VIDING FOR THE
X Furt	her documents are listed in the continuation of box C.	Patent family members are listed in annex.
"A" docume consider filling of the which citatio	ent which may throw doubts on priority claim(s) or is cited to establish the publication date of another in or other special reason (as specified) send to the form of the referring to an oral disclosure, use, exhibition or	"T" later document published after the International filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention  "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone  "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such document is
other	means ent published prior to the international filing date but han the priority date claimed	ments, such combination being obvious to a person skilled in the art.  *8.* document member of the same patent family
	actual completion of the international search	Date of mailing of the international search report
2	5 February 2000	03/03/2000
Name and	mailing address of the ISA  European Patent Office, P.B. 5818 Patentiaan 2  NL - 2280 HV Rilswijk	Authorized officer
	Tel. (+31–70) 340–2040, Tx. 31 651 epo nl, Fax: (+31–70) 340–3016	Chambonnet, F

1

### INTERNATIONAL SEARCH REPORT

rational Application No PCT/US 99/20675

	tion) DOCUMENTS CONSIDERED TO BE RELEVANT  Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
ategory *	Citation of document, with indication, where appropriate, or an arrangement of the citation of document, with indication, where appropriate, or an arrangement of the citation of document, with indication, where appropriate, or an arrangement of the citation of document, with indication, where appropriate, or an arrangement of the citation of document, with indication, where appropriate, and the citation of the	
	HASELHORST D ET AL: "STABLE PACKAGING CELL LINES AND HIV-1 BASED RETROVIRAL VECTOR SYSTEMS" GENE THERAPY,GB,MACMILLAN PRESS LTD.,	
	BASINGSTOKE, vol. 1, no. SUPPL. 02, 18 November 1994 (1994-11-18), page S14 XP002063698 ISSN: 0969-7128 the whole document	
•	ST LOUIS D ET AL: "CONSTRUCTION AND CHARACTERIZATION OF HIV-1 RETROVIRAL VECTORS AND REPLICATION-DEFECTIVE HIV-1 PACKAGING CELL LINES" INTERNATIONAL CONFERENCE ON AIDS AND THE STD WORLD CONGRESS, XX, XX, 1 June 1993 (1993-06-01), page 244 XP002063695 the whole document	1
A	CARROLL R ET AL: "A HUMAN IMMUNODEFICIENCY VIRUS TYPE 1 (HIV-1)-BASED RETROVIRAL VECTOR SYSTEM UTILIZING STABLE HIV-1 PACKAGING CELL	1
	LINES" JOURNAL OF VIROLOGY,US,THE AMERICAN SOCIETY FOR MICROBIOLOGY, vol. 68, no. 9, 1 September 1994 (1994-09-01), pages 6047-6051, XP002063697 ISSN: 0022-538X the whole document	
X	HOLLER T P ET AL: "HIV1 INTEGRASE EXPRESSED IN ESCHERICHIA COLI FROM A SYNTHETIC GENE" GENE,NL,ELSEVIER BIOMEDICAL PRESS. AMSTERDAM, vol. 136, 22 December 1993 (1993-12-22), pages 323-328, XP000199775 ISSN: 0378-1119	39
A	the whole document  ANDRE S ET AL: "INCREASED IMMUNE RESPONSE ELICITED BY DNA VACCINATION WITH A SYNTHETIC GP120 SEQUENCE WITH OPTIMIZED CODON USAGE"  JOURNAL OF VIROLOGY, US, THE AMERICAN SOCIETY FOR MICROBIOLOGY, vol. 72, no. 2, 1 February 1998 (1998-02-01), pages 1497-1503, XP002073767  ISSN: 0022-538X	39-43

1